MOLECULAR AND PHENOTYPIC ANALYSIS OF *SALMONELLA* BIOFILM FORMATION: EXPLORING THE LINKS BETWEEN SURVIVAL, VIRULENCE, AND TRANSMISSION

A Thesis Submitted to the College of
Graduate Studies and Research
In Partial Fulfillment of the Requirements
For the Degree of Doctor of Philosophy
In the Department of Microbiology and Immunology
University of Saskatchewan
Saskatoon

By

KEITH DAVID MACKENZIE

PERMISSION TO USE

In presenting this thesis/dissertation in partial fulfillment of the requirements for a Postgraduate degree from the University of Saskatchewan, I agree that the Libraries of this University may make it freely available for inspection. I further agree that permission for copying of this thesis/dissertation in any manner, in whole or in part, for scholarly purposes may be granted by the professor or professors who supervised my thesis/dissertation work or, in their absence, by the Head of the Department or the Dean of the College in which my thesis work was done. It is understood that any copying or publication or use of this thesis/dissertation or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of Saskatchewan in any scholarly use which may be made of any material in my thesis/dissertation.

Requests for permission to copy or to make other uses of materials in this thesis/dissertation in whole or part should be addressed to:

Head of the Department of Microbiology and Immunology
107 Wiggins Rd.
University of Saskatchewan
Saskatoon, Saskatchewan
S7N 5E5
Canada

OR

Dean
College of Graduate Studies and Research
University of Saskatchewan
107 Administration Place
Saskatoon, Saskatchewan S7N 5A2
Canada

ABSTRACT

Pathogenic *Salmonella* strains are responsible for millions of human and livestock infections each year. The mechanisms of *Salmonella* pathogenesis are of great interest, along with the capacity of strains to survive in the environment and complete the transmission cycle. This survival is predicted to be related to a specific physiology called a biofilm. Biofilms are communities of cells within a self-produced extracellular matrix that are often associated with a physical surface. For *Salmonella*, the biofilm phenotype is activated by the transcriptional regulator CsgD and is associated with the production of an extracellular matrix consisting of protein polymers and exopolysaccharides. *Salmonella* biofilm formation is induced during growth at low temperatures and in conditions of nutrient limitation and low osmolarity. The biofilm phenotype is highly conserved across nontyphoidal *Salmonella* strains that briefly colonize the host and cause gastroenteritis. It is hypothesized that biofilm formation is important for increasing the transmission success of nontyphoidal *Salmonella* by enhancing their persistence in non-host environments.

Salmonella biofilms have traditionally been studied as a population-level phenotype associated with colony formation, known as the red, dry, and rough (rdar) morphotype. However, *Salmonella* grown in liquid broth cultures under biofilm-inducing conditions form clonal subpopulations of multicellular aggregates and planktonic cells. This phenomenon is attributed to bistable expression of CsgD, where aggregated cells exist in a CsgD-ON state and planktonic cells are associated with a CsgD-OFF state. We performed comparative transcriptomic sequencing (RNA-seq), which revealed 1856 genes that were differentially expressed between these two *S.* Typhimurium cell subpopulations. Multicellular aggregates were associated with increased gene expression typical of *Salmonella* biofilm formation, including nutrient scavenging, reactive oxygen species defenses, and osmoprotection. In contrast, planktonic cells were associated with higher expression of multiple virulence pathways associated with the SPI-1 and SPI-2 type three secretion systems, cell motility, and chemotaxis. Increased synthesis of the SPI-1 type three secretion system in planktonic cells correlated with enhanced invasion of polarized Caco-2 human intestinal cells.

We modified an existing Tn7-based transposition system to generate chromosomally marked strains of *Salmonella* to facilitate tracking of multicellular aggregates and planktonic cells in competitive fitness assays. Planktonic cells were associated with increased virulence in mice compared to multicellular aggregates. However, when these same cell subpopulations were exposed to desiccation, multicellular aggregates were associated with greater cell survival and the virulence advantage of planktonic cells was lost. We hypothesize that bistable CsgD expression and the generation of specialized cell types may represent a form of bet hedging, where planktonic cells are adapted for direct host-to-host transmission, and multicellular aggregates can survive long-term in the environment to cause infections later. This strategy would prepare nontyphoidal *Salmonella* for the unpredictable nature of the fecal-oral transmission process and improve their potential to cause future infections.

Salmonella serovars that cause systemic disease within a restricted range of hosts have been shown to be biofilm-negative. In sub-Saharan Africa, a phylogenetically distinct group of nontyphoidal *Salmonella* has recently been identified for its role in an emerging epidemic of invasive extraintestinal infections. These invasive nontyphoidal *Salmonella* are associated with chronic persistence within the human host and do not have an identified environmental reservoir. We compared the biofilm phenotype of two invasive nontyphoidal *Salmonella* strains (*S.* Typhimurium D23580 and *S.* Enteritidis D7795) to a panel of strains consisting of 'typical' gastroenteritis-causing, nontyphoidal *Salmonella* and *Salmonella* strains that cause systemic typhoid fever. Both strains of invasive nontyphoidal *Salmonella* demonstrated an impaired biofilm phenotype, which we attributed to strain-specific genetic polymorphisms. We predict that the impaired biofilm phenotype of invasive nontyphoidal *Salmonella* correlates with their occupation of the systemic niche within the host and a reduced capacity to survive in the environment.

My research has brought insight into how pathogenic *Salmonella* strains are able to navigate through unpredictable areas of their lifecycle and increased our understanding of their potential transmission mechanisms.

ACKNOWLEDGEMENTS

I am forever grateful for the incredible group of family, friends, and mentors that I have met along the route of life and science. Aaron – thank you for seeing my potential and bringing out the best in me. You trusted me with the task of helping build the foundation of your laboratory and invested an incredible amount of time and effort into my education. Your insight has brought balance not only to my perspective of science, but also to my perspective of life. Thank you for being a mentor, colleague, and friend. I extend my gratitude to the incredible group of individuals that I have had the pleasure of working with during my time in the White laboratory. Dylan, Sumudu, and Yejun, I would not have made if through the marathon-style workdays without your support. Thank you for making me smile when it felt as though smiling was impossible. Leia, Cynthia, Nancy, Melissa, Michael, Nikole, and Dakoda – this work would not have been possible without you. Thank you for bringing an incredible dynamic to the laboratory. Wolfgang and Shirley, thank you for adopting me into your laboratory and providing unparalleled support. Over my years in graduate school I have become close friends with an incredible group of contemporaries. Amanda, Adam, Patricia, Brett, Tara, Ellen, Jean, Kristen, Ryan, Landon, Glenn, Kim, and Alyssa – thank you for being my family during this academic pursuit. You have walked this path along with me and have brought your own perspectives to how one balances science with sanity. Thank you for showing me the future is bright.

I belong to an incredibly strong family that has shaped me into who I am today. Mom and dad, thank you for your wisdom, patience, and support. Thank you for picking me up when I faltered. Thank you for accepting me for who I am, and for teaching me how to love myself. I would not have achieved this moment without you. Heather, thank you for being an incredible role model and the big sister that everyone hopes they could have. Neil, thank you for teaching me how to work hard and play hard. Jill, I am grateful that you were there to traverse the Ph.D. valley with me. Thank you for inspiring and encouraging me every day. I am also grateful to the friends I have made during my time at the University of Saskatchewan. Lexie, Trent, Shantelle, Aimee, Adam, Michelle, and Dielle – thank you for sharing your lives with me. You have made this city my home

away from home. Thank you to Lisa and Erin for being there for me every day. Alex, you have been an incredible friend. Thank you for being a friendly face during hard days, long nights, and working weekends.

Len, you are my heart. You came into my life during the hardest part of my graduate studies. I would not have made it through this without you. You have shown me that there is more to this life, and you challenged me to become something much more than I ever imagined I could be. Thank you for your strength, humour, patience, and companionship. I am forever grateful to share this life with you.

DEDICATION

I dedicate this thesis to my partner, Len.
Pablo Picasso once said, "It took me four years to paint like Raphael, but a lifetime to paint like a child."
Thank you for reminding me how to colour.

This thesis is also dedicated to my mom, Elizabeth.
Thank you for being there every day and for involuntarily signing up for two doctoral degrees.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF APPENDICES

LIST OF ABBREVIATIONS

ACP	Acyl carrier protein
adr	AgfD-regulated
AMP	Adenine monophosphate
ANOVA	Analysis of variance
ATPase	Adenosine triphosphatase
bap	Biofilm-associated protein
bp	Base pairs
bcs	Bacterial cellulose biosynthesis
c-di-GMP	Bis-(3'-5')-cyclic dimeric guanosine monophosphate
CFU	Colony forming units
CI	Competitive index OR confidence interval
Cm	Chloramphenicol
cpx	Conjugative plasmid expression
crl	Curli
CPS	Counts per second
csg	Curli synthesis gene
DGC	Diguanylate cyclase
DMEM	Dulbecco's Modified Eagle Medium
DNA	Deoxyribonucleic acid
env	Envelope
FACS	Fluorescence-assisted cell sorting
FDR	False discovery rate
FLP	Flippase
FUN	Function unknown
GALT	Gut-associated lymphoid tissue
gcv	Glycine cleavage
GFP	Green fluorescent protein
GI	Gastrointestinal
GMP	Guanosine monophosphate

GTP	Guanosine triphosphate
GTPase	Guanosine triphosphatase
HCM	Host cell membrane
hfq	Host factor for phage Q beta
HILIC	hydrophilic interaction
HIV	Human immunodeficiency virus
H-NS	Histone-like nucleoid structuring
IHF	Integration host factor
IL	Interleukin
IM	Inner membrane
iNTS	Invasive nontyphoidal *Salmonella*
IPTG	Isopropyl β-D-1-thiogalactopyranoside
IVET	*In vivo* expression technology
kb	Kilobase
kDa	Kilodaltons
KEGG	Kyoto Encyclopedia of Genes and Genomes
Kn	Kanamycin
lac	Lactose
M cell	Microfold cell
MLN	Mesenteric lymph nodes
mlr	MerR-like regulator
MOI	Multiplicity of infection
MRM	Multiple-reaction monitoring
OM	Outer membrane
omp	Outer membrane protein
NSERC	National Sciences and Engineering Council
NTS	Nontyphoidal *Salmonella*
PAMP	Pathogen-associated molecular pattern
PBS	Phosphate buffered saline
PDE	Phosphodiesterase
pGpG	5’-phophoguanyl-(3’-5’)-guanosine

pH	Potential of hydrogen
pil	Pili
rdar	Red, dry, and rough
RFU	Relative fluorescent units
RNA	Ribonucleic acid
RPKM	Reads per kilobase million
rpo	RNA polymerase
SCV	*Salmonella*-containing vacuole
SPI-1	*Salmonella* pathogenicity island
sRNA	Small ribonucleic acid
STM	*Salmonella* Typhimurium
T3SS	Type three secretion system
TCA	Tricarboxylic acid
tns	Transposase
TP	Time point
UV	Ultraviolet

1.0 LITERATURE REVIEW

1.1 *Salmonella*

1.1.1 Classification

1.1.1.2 Taxonomy and Nomenclature

Salmonella is named after Dr. David E. Salmon, who first isolated the bacterium from porcine intestine in 1884 (1). Salmonellae are gram-negative, motile, facultative anaerobic rods that measure 0.7-1.5 μm in width by 2.0-5.0 μm in length (2). *Salmonella* is a genus belonging to the kingdom Eubacteria, class Gammaproteobacteria, order Enterobacteriales, and family Enterobacteriaceae (Fig. 1.1). The genus is further subdivided into species, subspecies, serovars, and strains based on a mixture of taxonomic and serological classifications. A newly identified strain of *Salmonella* is categorized into species and subspecies according to deoxyribonucleic acid (DNA) relatedness at the genomic level, originally shown through DNA-DNA hybridization, and the presence or absence of 11 biochemical traits (3, 4). Two species have been defined, *Salmonella bongori* and *Salmonella enterica,* of which *Salmonella enterica* is further split into six subspecies that are designated by a Roman numeral and name (I, *enterica*; II, *salamae*; IIIa, *arizonae*, IIIb, *diarizonae*, IV, *houtenae*; and VI, *indica*) (5). Serological characterization, originally established by Kauffman and White, was historically used to classify newly isolated *Salmonella* strains and continues to be used in addition to traditional taxonomic classification (6). *Salmonella* isolates are assigned to a serovar based on the presented combination of flagellar antigens (H1 and H2) and (lipopolysaccharide) oligosaccharide (O) or capsular polysaccharide (K) antigen (7). Using this system, more than 2600 serovars have been identified, with their given name reflecting their combination of antigens, or in the case of serovars in subspecies *enterica*, a name representing their associated disease, host specificity, geographic origin, or relationship to other identified serovars (1, 7, 8). *Salmonella enterica* subspecies *enterica* is the most well-represented amongst serovars and disease, accounting for approximately 60% of all serovars identified and greater than 95% of *Salmonella* isolates obtained from

humans and domestic mammals (9, 10). In contrast, *Salmonella* isolates belonging to the remaining species and subspecies are normally obtained from cold-blooded hosts, but are sometimes able to cause infections in humans (11).

In presenting a *Salmonella enterica* strain, the genus, species, and subspecies names are provided, followed by "serovar" and the serovar name, which is capitalized and non-italicized (7). For all subsequent mentions, the name is shortened and only the genus and serovar name are given (e.g., *Salmonella* ser. Typhimurium or *S.* Typhimurium in place of *Salmonella enterica* subspecies *enterica* serovar Typhimurium) (1, 7). For the purposes of my thesis, I will only be talking about isolates of *S. enterica* subspecies *enterica* and will refer to them by their common serovar names.

1.1.1.2 Classification by Disease and Host Range

Aside from taxonomic convention, *Salmonella* isolates are often described based on their associated disease or for their associated range of susceptible hosts. Serovars associated with disease in humans are divided into nontyphoidal Salmonellae (NTS), which classically induce a self-limited gastroenteritis, or typhoidal Salmonellae, the causative agents of enteric fever (Fig. 1.1). However, recent increases in the global cases of nontyphoidal *Salmonella* associated with and invasive, systemic disease has resulted in an addendum to this reference system. Strains of nontyphoidal *Salmonella* associated with invasive disease are as such referred to as invasive nontyphoidal Salmonellae, or iNTS. Serovars may also be described in the literature according to their host range. In this system, a *Salmonella* serovar may be classified as host-generalist, host-adapted, or host-restricted (12). Host generalist *Salmonella* serovars, such as *S.* Typhimurium and *S.* Enteritidis, are able to colonize multiple host species (12). In contrast, host-adapted and host-restricted serovars are associated with an invasive extra-intestinal disease in the host. Host-adapted species, including *S.* Choleraesuis and *S.* Dublin, have narrowed their host range to preferentially infect a particular host (in these examples, swine or cattle), but retain the capacity to cause infections in other hosts (12). In contrast, host-restricted serovars, such as *S.* Typhi or *S.* Paratyphi in humans or *S.* Gallinarum in fowl, produce a systemic disease in only one species (12). As the majority of *Salmonella* research is

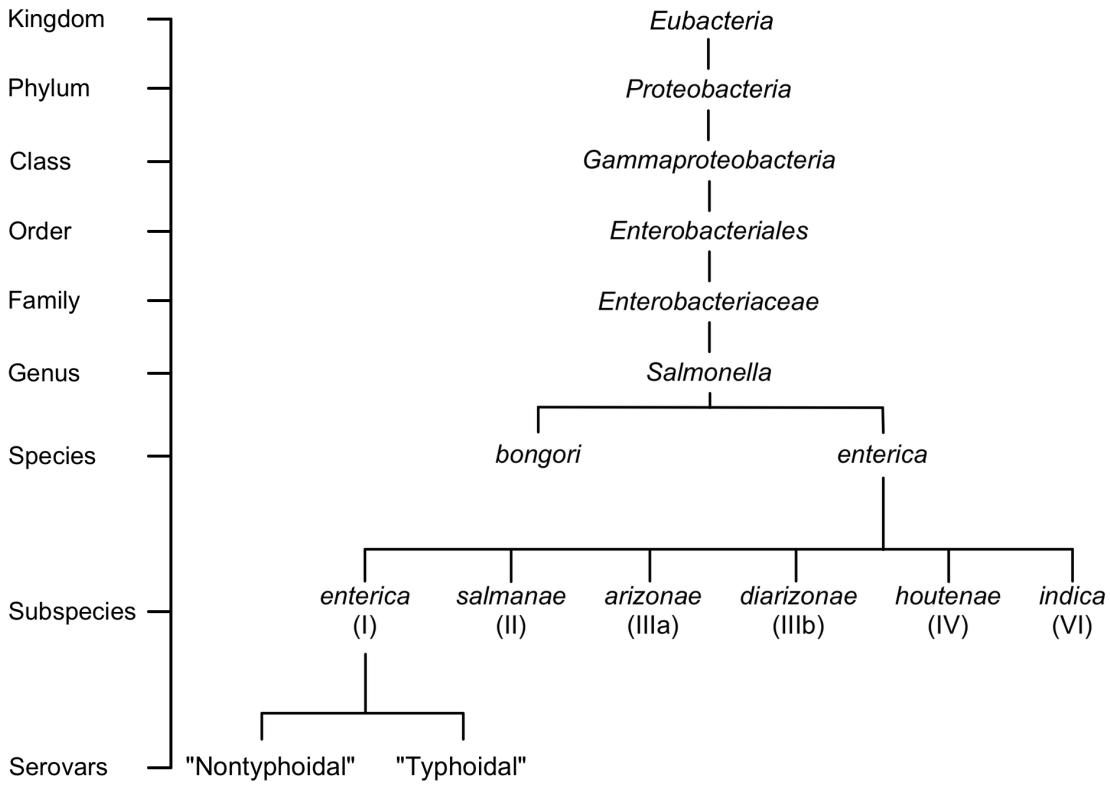

Figure 1.1. *Salmonella* taxonomy and general classifications.

The genus *Salmonella* is classified into species, subspecies, and serovars based on the White-Kauffman-Le Minor scheme. Serovars are often grouped into nontyphoidal or typhoidal categories; however, this referencing approach is not a part of the official *Salmonella* classification scheme.

focused on disease, pathogenesis, and epidemiology affecting humans, I will refer to typhoidal and nontyphoidal *Salmonella* classifications.

1.1.2 Nontyphoidal, typhoidal, and invasive nontyphoidal *Salmonella*

1.1.2.1 Epidemiology

Nontyphoidal *Salmonella* serovars are responsible for the majority of global cases of human salmonellosis. It is estimated that nontyphoidal *Salmonella* account for approximately 153 million annual cases of gastroenteritis, with 57 000 deaths (13). Nontyphoidal *Salmonella* infections are associated with the ingestion of contaminated food or water, but can also be transmitted directly between people or from zoonotic sources, such as domestic or food animals, through the fecal-oral route (10). Outbreaks of nontyphoidal *Salmonella* have traditionally been associated with a wide range of food products, including animal-based (meat, poultry, eggs), plant-based (tomatoes, sprouts, melons, lettuce, mangoes, raw almonds) and processed foods (powdered infant formula, dry seasonings, cereals, peanut butter) (10). *S.* Enteritidis and *S.* Typhimurium are the serovars most frequently associated with nontyphoidal *Salmonella* infections (14, 15). On a global scale, *S.* Enteritidis was responsible for 65% of all human infections, followed by *S.* Typhimurium with 12% and *S.* Newport with 4% (14). For North America, nontyphoidal *Salmonella* illnesses are more frequently caused by *S.* Typhimurium (29% of all human infections) than by *S.* Enteritidis (21%) (14).

Typhoidal *Salmonella* infections are estimated to cause approximately 21 million illnesses and 145 000 deaths worldwide (13). Typhoidal *Salmonella* infections are unequally distributed between high-income and low-income nations. In the United States and some European countries, the annual rate of typhoidal *Salmonella* infections is less than 10 infections per 100 000 population (20). However, many Asian countries (China, Indonesia, and Vietnam) have greater than 100 cases per 100 000 population, with Pakistan and India have a disproportionately high incidence of typhoidal infection (451.7 and 214.2 cases per 100 000, respectively) (21). Typhoidal *Salmonella* serovars are transmitted primarily from person to person through food and water contaminated with human feces; as such, infections are more frequent in low-income countries that lack

available safe water resources and have poor sanitation standards (17). Cases of typhoidal infections in high-income countries are usually the result of patients travelling to endemic areas, but can also be spread by individuals that are chronically infected with *S.* Typhi (17).

The number of incidences of nontyphoidal *Salmonella* infections associated with an invasive disease has increased dramatically in recent years, resulting in an estimated 3.4 million illnesses and 681 000 deaths annually (16). The continent of Africa accounted for almost 2 million cases of iNTS infections alone in 2010, representing more than half of the total number of global cases of this disease (16). Nontyphoidal *Salmonella* strains associated with an invasive disease are speculated to have a greater dependency on direct transmission from person to person (16). Immunosuppression is an important host factor associated with the spread of iNTS disease. These infections occur predominately in children between 6 and 18 months of age and in adults between 25 and 40 years old (17). For children, the predominant host risk factors for iNTS disease are human immunodeficiency virus (HIV) infection, malnutrition, and malaria, while advanced HIV infection is the main risk factor for adults (18). Further, cases of iNTS disease in children and adults are strongly correlated with the rainy season in sub-Saharan Africa, which could be the result of waterborne transmission, malaria, or malnutrition during this season (17, 19). Community-acquired iNTS infections in Africa are most frequently associated with *S.* Typhimurium, accounting for 65.2% of all iNTS infections, or *S.* Enteritidis, accounting for 33.1%, (15).

1.1.2.2 Disease Symptoms in Humans

Nontyphoidal *Salmonella* infections in immunocompetent individuals normally result in a severe, self-limiting gastroenteritis (22-24). The characteristic symptoms last for 4 to 7 days and include profuse amounts of non-bloody diarrhea, as well as nausea, vomiting, and abdominal cramps (18, 23, 24). Symptoms are preceded by an incubation period that averages 6 to 12 hours, but can be up to 72 hours post-infection (22-24). Continued shedding of *Salmonella* in the feces occurs long after symptoms have cleared; with a median shedding time of 1 month post-symptoms for adults, or 7 weeks for children under the age of 5 years (22). Antibiotic use is normally discouraged during such

infections in immunocompetent patients, as some studies have demonstrated increased duration of asymptomatic shedding of the pathogen following treatment (22, 23). In most nontyphoidal *Salmonella* infections, the pathogen is spatially limited to the gastrointestinal tract (24). However a secondary bacteremia can arise in approximately 5% of cases, where the pathogen has disseminated past the lamina propria and invaded further into the body (22-24). In such cases, treatment with bactericidal antibiotics is recommended (third generation cephalosporin or intravenous fluoroquinolone) and results in clearance of the pathogen after 7 to 14 days of therapy (23).

Typhoidal *Salmonella* infections in humans are caused by *S.* Typhi and *S.* Paratyphi and are associated with enteric fever. The incubation period preceding symptoms is longer for typhoidal *Salmonella* infections, averaging 1 to 2 weeks but extending for as long as 60 days is some cases (25). As the pathogen enters the bloodstream, patients present with the initial symptoms of headache, gradual fever, and lethargy (25). This fever can persist for up to 4 weeks if left untreated (23). Additional symptoms are infrequent in their presence between cases, but may include the appearance of "rose spots" with a 2 to 4 mm diameter on the abdomen and chest, bradycardia, enlargement of the spleen and liver, anorexia, vomiting, and a dry, coated tongue (25). In contrast to gastroenteritis caused by nontyphoidal *Salmonella*, these infections can result in constipation for immunocompetent adults, but diarrhea may occur in young children or in adults with HIV infection (25). Typhoidal *Salmonella* infections are usually successfully treated with 5 to 7 days of antimicrobial treatment with fluoroquinolones such as ciprofloxacin (23). However, the rise of typhoidal *Salmonella* strains resistant to quinolones in some parts of the world, such as Asia, necessitate the use of alternative drugs such as ceftriaxone, cefixime, or azithromycin (23). Left untreated, 10 to 15% of cases result in complications that may involve gastrointestinal bleeding, intestinal perforation, and encephalopathy (25). In an additional 5 to 10% of cases, the patient will experience a relapse in infection 2 to 3 weeks following fever (25). The period of asymptomatic carriage during this complication of typhoidal *Salmonella* infection is generally longer than that for nontyphoidal *Salmonella*, with 10% of patients shedding the pathogen for 3 months, and 1 to 4% of patients shedding for greater than 1 year (25).

Invasive nontyphoidal *Salmonella* strains cause a typhoid-like illness and are often associated with immunocompromised individuals (17, 18). These infections are mainly characterized by a primary bacteremia, and symptoms of diarrhea are observed in only 20 to 50% of cases (18). Patients frequently present with respiratory symptoms (i.e., an increased respiratory rate and chest crackles) which are often due to co-infection with other pathogens, such as *Mycobacterium tuberculosis* or *Streptococcus pneumoniae* (17, 18). Additional symptoms, such as hepatomegaly and splenomegaly, are less frequently observed (18); however, splenomegaly was reported as a useful indicator of iNTS disease in adults with HIV infection (26). Invasive nontyphoidal *Salmonella* infections can further develop into bacterial meningitis, which is associated with a high case fatality in both children (52%) and adults (80%) (18). Further, approximately 20 to 30% of HIV-infected adults that have received treatment for iNTS disease will experience a relapse in infection due to an identical *Salmonella* strain between 4 and 6 months after the primary illness, suggesting that the pathogen may persist within the patient (17). The most striking feature of iNTS disease is its high case fatality despite correct microbiological diagnosis and treatment – in Africa, up to 22% of children and 47% of adults succumb to the infection despite appropriate intervention (27). To further complicate efficacious treatment, some iNTS strains are associated with multidrug resistance, namely ampicillin, chloramphenicol, and co-trimoxazole, resulting in the need to use expensive antimicrobial treatments such as third-generation cephalosporins and fluoroquinolones (18). For HIV-infected individuals, initiation of antiretroviral therapy has been reported as effective in preventing relapse in iNTS infections (18).

1.1.2.3 Pathogenesis of Disease

Central to *Salmonella* pathogenesis is its ability to modify host cell biology via two type-three secretion systems (T3SS), T3SS-1 and T3SS-2 (28). These specialized organelles span the bacterial inner and outer membranes and allow for the delivery of effector proteins into the cytoplasm of eukaryotic cells (29, 30). Type three secretion systems are found exclusively in gram-negative bacteria (28, 31). For *Salmonella*, genes for the T3SS-1 or T3SS-2 apparatus, regulatory components, and nearly all associated effector proteins are found within horizontally acquired DNA regions known as

Salmonella Pathogenicity Islands (SPIs) (30). The SPI-1 region contains genes associated with T3SS-1, is found in all serovars of *S. enterica* and *S. bongori*, and is important for the invasion of intestinal epithelial cells (32, 33). In contrast, the full-length SPI-2 region harbouring genes for T3SS-2 is present exclusively in *S. enterica*, and is associated with the intracellular survival of *Salmonella* within eukaryotic cells (33).

Following *Salmonella* entry into the host via contaminated food or water, cells travel through the digestive system and localize to the distal ileum and colon of the gastrointestinal tract (34). *Salmonella* cells rely on flagellar motility and chemotaxis systems to traverse the intestinal mucus layer and identify sites on the host cell that are permissive for invasion (35). Fimbriae and other protein adhesins on the bacterial cell surface initiate association of the pathogen to the targeted epithelial cell; however, the needle complex of the SPI-1 T3SS is critical for stabilizing this interaction (36, 37). It is currently hypothesized that expression and synthesis of the SPI-1 T3SS is induced by several important cues provided by the local host environment, including low oxygen tension, high osmolarity, near-neutral pH, and acetate production levels from the resident microflora (38). The secretion apparatus is established in a step-wise manner, requiring formation of the basal body within the bacterial cell membranes to facilitate secretion of the needle complex (39). Proteins attached to the end of the SPI-1 T3SS needle, collectively referred to as the translocon, are then inserted into the host cell membrane, creating a pore that allows for the injection of *Salmonella* effector proteins into the host cell (36).

Establishment of a secretion-competent SPI-1 T3SS is a mandatory prerequisite for the process of host cell invasion (39). Of the repertoire of 11 effector proteins known to be secreted by the SPI-1 T3SS, early translocation of the proteins SopE, SopE2, SopB, and SipA is essential to promoting bacterial-mediated endocytosis (reviewed in (40)). SopE, SopE2, and SopB act in a functionally redundant manner by inducing the host cell's Rho GTPases, which activate the signaling transduction pathway that leads to actin cytoskeleton rearrangement. The unique functional activities of the effector protein SipA and the translocon protein SipC coordinate the localization of this cytoskeleton rearrangement to the site of interaction between the bacterial and host cell surfaces, resulting in dramatic actin filament accumulation and host cell membrane ruffling.

Salmonella cells are then internalized into the host within endosomes termed *Salmonella*-containing vacuoles (SCVs), a process facilitated by SopB and an additional effector protein, SopD. Following bacterial uptake, the host cell cytoskeleton is then returned to a state of homeostasis through deactivation of the Rho GTPases by another T3SS effector, SptP.

Following the invasion process, *Salmonella* depend on the SPI-2 T3SS and associated effectors to further manipulate the host cell environment and promote pathogen survival and replication within the SCV. Expression and activation of the SPI-2 T3SS is induced by changes in the SCV environment, including vacuolar acidification (pH 5 to 5.5), decreased concentrations of divalent cations and phosphates, and the presence of antimicrobial peptides (41). Many of the effector proteins secreted by the SPI-2 T3SS contribute important functions for manipulating the movement of the SCV along host cell microtubules. For example, the effector proteins SopB and SsaB prevent bacterial killing by inhibiting SCV trafficking and fusion to lysosomes (42). A series of additional effector proteins (SifA, SseF, SseG, and PipB2) are responsible for the redirection of the SCV towards the Golgi apparatus, which may be an important location for the acquisition of nutrients and cell components from other host cell vesicles (30, 40). Following this localization near the Golgi apparatus, two effector-mediated processes are activated during *Salmonella* cell replication in the SCV. To maintain the membrane integrity of the SCV during this cell replication, interactions between the effectors SspH2 and SseI and host proteins filamin and profilin result in the polymerization of an actin meshwork that surrounds the vacuolar membrane (30, 40). Further, *in vitro* studies have shown that *Salmonella* replication coincides with the production of *Salmonella*-induced filaments that extend from the SCV. Several SPI-2 T3SS effectors, including SifA, SseJ, and SopD2, are implemented in inducing the formation of these SIFs, which are hypothesized to provide a potential source of nutrients during pathogen replication (40).

Pathogen cells that traverse the epithelial cell layer encounter tissue mononuclear cells (i.e., macrophages and dendritic cells) within the lamina propria, resulting in the uptake of the pathogen into a phagosome (34). Detection of pathogen-associated molecular patterns (PAMPs) and components injected by the pathogen during its uptake into host cells elicits the production of a proinflammatory immune response (34). This

response inhibits the spread of nontyphoidal *Salmonella* past the lamina propria in three ways: by activating infected macrophages and inducing killing of intracellular *Salmonella*, through recruitment of neutrophils to the infection site for extracellular killing, and by stimulating epithelial cells to release antimicrobial peptides into the intestinal lumen to control the replication of nontyphoidal *Salmonella* cells (34). While host inflammation effectively controls those nontyphoidal *Salmonella* that have invaded the epithelial cell layer, it acts as a potent stimulator of growth of the nontyphoidal *Salmonella* population in the intestinal lumen (34). Recent studies have shown two mechanisms by which nontyphoidal *Salmonella* are able to exploit the host inflammatory response. Epithelial cells release the antimicrobial agent lipocalin-2, a molecule that binds to enterochelin, an iron chelation molecule used by gram-negative bacteria in the gut to acquire iron from the host (43). In addition to enterochelin, nontyphoidal *Salmonella* are able to produce a second iron chelation molecule, salmochelin, which cannot be bound by lipocalin-2 (43). As a result, nontyphoidal *Salmonella* cells in the intestinal lumen continue to replicate while the local microbiota are starved for iron. The low-oxygen conditions of the intestinal lumen promote the establishment of an anaerobic microbiota that use fermentation to derive energy from available amino acids and complex polysaccharide (34). Hydrogen sulfide is produced as a byproduct of this fermentation, which is immediately converted to thiosulfate by the epithelial cell layer of the colon (44). During inflammation, neutrophils infiltrate the intestinal lumen and release reactive oxygen species molecules as part of the mechanism for the extracellular killing of bacterial pathogens (45). The association of reactive oxygen species with thiosulfate molecules results in tetrathionate, which can be used by nontyphoidal *Salmonella* as an electron acceptor during anaerobic respiration (44). In addition to activating a metabolic response that promotes growth of the pathogen, anaerobic respiration further allows nontyphoidal *Salmonella* to utilize carbon sources that would otherwise metabolize poorly during aerobic fermentation (46). Altogether, nontyphoidal *Salmonella* can use these mechanisms to promote their own growth at the expense of the existing host microbiota.

Two features that distinguish enteric fever from gastroenteritis are the absence of inflammation and an innate immune response, and the replication of typhoidal

Salmonella in the systemic compartment of the host. These changes in pathogenesis are linked to genomic differences between typhoidal and nontyphoidal *Salmonella*. Approximately 200 functional genes in nontyphoidal *Salmonella* have been inactivated or functionally disrupted in *S.* Typhi and *S*. Paratyphi A (47). Many of these mutations in *S.* Typhi affect processes used by nontyphoidal *Salmonella* to induce intestinal inflammation, including motility and chemotaxis, adherence to and invasion of host cells, and loss of virulence factors associated with intracellular replication (47). The loss of these functions suggest that *S.* Typhi may gain access to the systemic compartment through a mechanism distinct from active invasion of intestinal epithelial cells. Although this mechanism remains elusive, it is hypothesized that microfold (M) cells that sample the intestinal lumen actively take in *S.* Typhi cells and transfer them to macrophages and dendritic cells within the gut-associated lymphoid tissue (GALT) of the Peyer's patches, located in the small intestine (48). Typhoidal *Salmonella* are unique/distinct from nontyphoidal serovars in that they are able to persist in this intracellular niche without activating the immune response and infiltration of neutrophils that would otherwise restrict typhoidal infections (49). Within these immune cells, typhoidal *Salmonella* are shuttled to other sites in the body associated with the mononuclear phagocyte system (previously known as the reticuloendothelial system), taking residence in such places as the liver, spleen, mesenteric lymph nodes, bone marrow, as well as the gall bladder (50, 51). Typhoidal *Salmonella* are thought to transfer back into the duodenum via the biliary tract, resulting in shedding of typhoidal cells in the feces and potential transmission to a new host (50).

The lack of inflammation associated with typhoidal *Salmonella* infections suggests important differences in the pathogen's surface antigens or its interactions with host cells. For example, the potential down-regulation in *S.* Typhi flagellar expression reduces the presence of this PAMP and results in decreased inflammation (52). Other gene mutations identified in *S.* Typhi include regulatory elements affecting the O-antigen structure, which may limit exposure of this important PAMP to immune cells (53). Further, the typhoidal *Salmonella* genome has approximately 300 to 400 unique genes that are absent in nontyphoidal *Salmonella*. Of these additional accessory genes, the Vi capsule plays an important role in reducing the host inflammatory response to the

presence of *S.* Typhi cells. Production of the Vi capsule limits complement deposition on the surface of *S.* Typhi cells, masks surface antigens that would normally activate the host immune response, and provides resistance to phagocytic killing (48, 50, 54). Further, the Vi capsule has also been demonstrated to induce increases in production of the cytokine interleukin 10 (IL-10), an important anti-inflammatory molecule (55). Vi-negative mutants of *S.* Typhi were unable to cause enteric fever in human infection trials (56). However, the Vi capsule cannot solely account for differences between typhoidal and nontyphoidal *Salmonella* infections, as this capsule is not expressed by other typhoidal serovars (i.e., *S.* Paratyphi A and *S.* Sendai) (49). Therefore, it is likely that serovar-specific combinations of gene acquisition and gene loss is responsible for the ability of typhoidal *Salmonella* strains to evade the host immune response.

1.1.2.4 Pathogen Transmission

Microorganisms are under evolutionary pressure to define a niche that allows for their continued survival and replication. For pathogens, addressing this requirement means finding a balance between optimizing their fitness within a given host and maximizing their potential for transmission to new hosts (57). The diversity of available hosts species presents an opportunity for pathogens to exploit new, additional niches for their own replication (58). Pathogens that successfully adopt new niches and broaden their host range are known as host generalists (59). However, pathogens that overcome the barriers to infecting a new host species face attempts by the host to constrain their ability to survive and reproduce (60, 61). Host pressures can drive genetic changes within the pathogen, some of which may permit the pathogen to replicate in the new niche at the cost of impacting its fitness in other host species (58, 59). These host-restricted pathogens face the new challenge of ensuring their transmission success to a narrowed pool of susceptible hosts (58).

Host generalist pathogens transmit between several potential host species. An estimated 60% of human pathogens and 80% of pathogens infecting domestic animals are zoonotic in nature (62, 63). Nontyphoidal *Salmonella* species represent host generalist pathogens that have adopted niches in mammals, birds, and reptiles (58). In contrast, some typhoidal *Salmonella* species demonstrate host restriction, establishing long-term

relationships in humans (*S.* Typhi and *S.* Paratyphi serovars) or chickens (*S.* Gallinarum, *S.* Pullorum) (12, 64). Other typhoidal *Salmonella* are host adapted, having specific preference for replication in certain species (*S.* Dublin and cattle, *S.* Choleraesuis and pigs) while still retaining the ability to infect humans (12). Despite this change in host range, nontyphoidal and typhoidal *Salmonella* serovars maintain a genetic relatedness at the species level (49). Therefore, *Salmonella* pathogens present an opportunity to study the biological factors that are important to different modes of transmission (58).

Typhoidal *Salmonella* depend upon establishing chronic persistence within their current host to increase the opportunities for subsequent transmission events. Between 5 and 10% of patients recovering from enteric fever experience a relapse in infection with the same typhoidal *Salmonella* strain, resulting in milder symptoms than before and fecal shedding of the pathogen for 3 weeks to 3 months following the initial infection (25, 50). While the mechanism behind this short-term persistence in the host is unclear, it is hypothesized that typhoidal *Salmonella* remain dormant within immature immune cells in the bone marrow (48). In contrast, between 2 and 4% of people living in areas endemic for enteric fever are associated with a chronic carrier state that involves asymptomatic carriage of typhoidal *Salmonella* for more than a year (48). Epidemiological studies hoping to identify factors associated with the chronic carrier state are difficult due to the asymptomatic nature of infections in these hosts (50). However, recent evidence points to persistence of typhoidal *Salmonella* within the gall bladder. It is currently hypothesized that typhoidal *Salmonella* cells first localize to the liver and replicate in the resident macrophage (Kupffer) cells before traveling to the gall bladder via the biliary tract (50). Bile, a digestive secretion with detergent and antimicrobial properties, contributes to the sterility of the gall bladder. In a landmark study assessing the incidence of gall bladder disease in patients associated with acute or chronic infections with typhoidal *Salmonella*, the authors detected gallstones in nearly 90% of chronically-infected patients (65). Research by the John Gunn and his colleagues revealed the presence of typhoidal *Salmonella* cells associated with gallstones. Microscopic analysis of this interaction suggested that typhoidal *Salmonella* cells use fimbrial protein structures on their surface to attach to gallstones, while growth in the presence of bile stimulates the production of protective extracellular polysaccharides (66). It is currently hypothesized that short-term

human carriers are mainly responsible for the transmission of enteric fever in endemic areas, while long-term chronic carriers are responsible for resurgence of infections in endemic regions despite attempts to control such infections (50).

Nearly all nontyphoidal *Salmonella* serovars associated with human disease demonstrate the ability to colonize multiple host species and induce gastroenteritis. This ability to infect a wide host range maximizes the potential for nontyphoidal *Salmonella* to directly transmit from one host to the next, via the fecal-oral route. However, a number of nontyphoidal *Salmonella* outbreaks in developed countries emphasize the importance of environmental reservoirs as an intermediary step for transmission of this pathogen. Several reported cases exemplify the ability of nontyphoidal *Salmonella* to persist in non-host environments. Surface runoff of water contaminated with animal feces was suspected as the source of two separate 2008 cases of drinking water contamination affecting communities in the United States (67, 68). Similarly, a study analyzing 288 cases of drinking water-related outbreaks in Canada between the years 1971 and 2001 noted water treatment practices and nearby wildlife as the most frequently reported sources of contamination (69). Several reports of nontyphoidal *Salmonella* outbreaks associated with fresh produce have traced contamination to irrigation water or animal manure used to fertilize fields (70-72). Of particular interest are two separate outbreaks of nontyphoidal *Salmonella* infections in the United States in 2002 and 2005, both of which were linked to a rare strain of *S.* Newport in tomatoes (70). In both outbreaks, investigators were able to trace back the unique strain to a Virginian farm, where the strain had been isolated from a contaminated pond used to irrigate the fields (70). The identification of this same rare strain in pond water samples taken years apart indicates the added importance of *Salmonella* persistence in environmental reservoirs (70). A similar case of nontyphoidal *S.* Typhimurium persistence was observed for a Danish pig farm associated with recurring infections in its herd (73). Samples collected by Baloda and colleagues over a two-year period revealed the presence of the same *Salmonella* clone in the piggery, the feed provided to the animals, and in the pig manure used to fertilize agricultural soil (73). Farmland soil treated with manure yielded viable *S.* Typhimurium cells for 14 days following spread, providing further evidence for the survival of nontyphoidal *Salmonella* within the non-host environment. The authors

hypothesized that the persistence of *Salmonella* in this setting could result in a cycle of re-infection of the pig herd from environmental reservoirs, potentially explaining the long-term presence of nontyphoidal *Salmonella* at the farm.

1.2 *Salmonella* Biofilms

Bacterial biofilms are communities of bacterial cells within a self-produced extracellular matrix that are often associated with a physical surface. For *Salmonella*, the biofilm phenotype is described as the rdar morphotype, named for the red, dry, and rough appearance of colonies grown on agar plates containing Congo Red dye (74, 75) (Fig. 1.2a). *Salmonella* biofilms in liquid cultures have also been described as pellicles, which refers to a film of cell growth that appears at the air-liquid interface (75-77) (Fig. 1.2b). The structure of the extracellular matrix consistently includes the presence of protein polymers (curli fimbriae) and extracellular polysaccharides (cellulose and O-antigen capsule) (74, 75, 78, 79). For *Salmonella*, biofilm formation is induced by conditions of nutrient starvation, low osmolarity, and growth at temperatures below 30˚C (74, 75, 80). The resulting biofilm phenotype has been demonstrated to enhance the persistence of *Salmonella* cells exposed to conditions of nutrient stress, desiccation, and disinfection (76, 80-84). *Salmonella* cells within a biofilm are able to cause infection; however, the biofilm phenotype is not a virulence adaptation (85). Biofilm formation is hypothesized to enhance *Salmonella* survival in the non-host environment and improve its odds of transmission to a susceptible host (85).

1.2.1 Structural and Regulatory Elements Responsible for Biofilm Formation

Curli fimbriae are protein polymers of repetitive subunits that are important for short-distance interactions between *Salmonella* cells within a biofilm (reviewed in (86)). The aggregative nature of cells expressing curli fimbriae was described in their original discovery within a strain of *E. coli* isolated from cattle manure (87). Similar fimbriae were later described in an enterotoxinogenic *Salmonella* Enteritidis strain isolated from a patient in India and given the name thin aggregative fimbriae to reflect their small structural diameter and aggregative nature (74). Collinson *et al.* demonstrated the unique

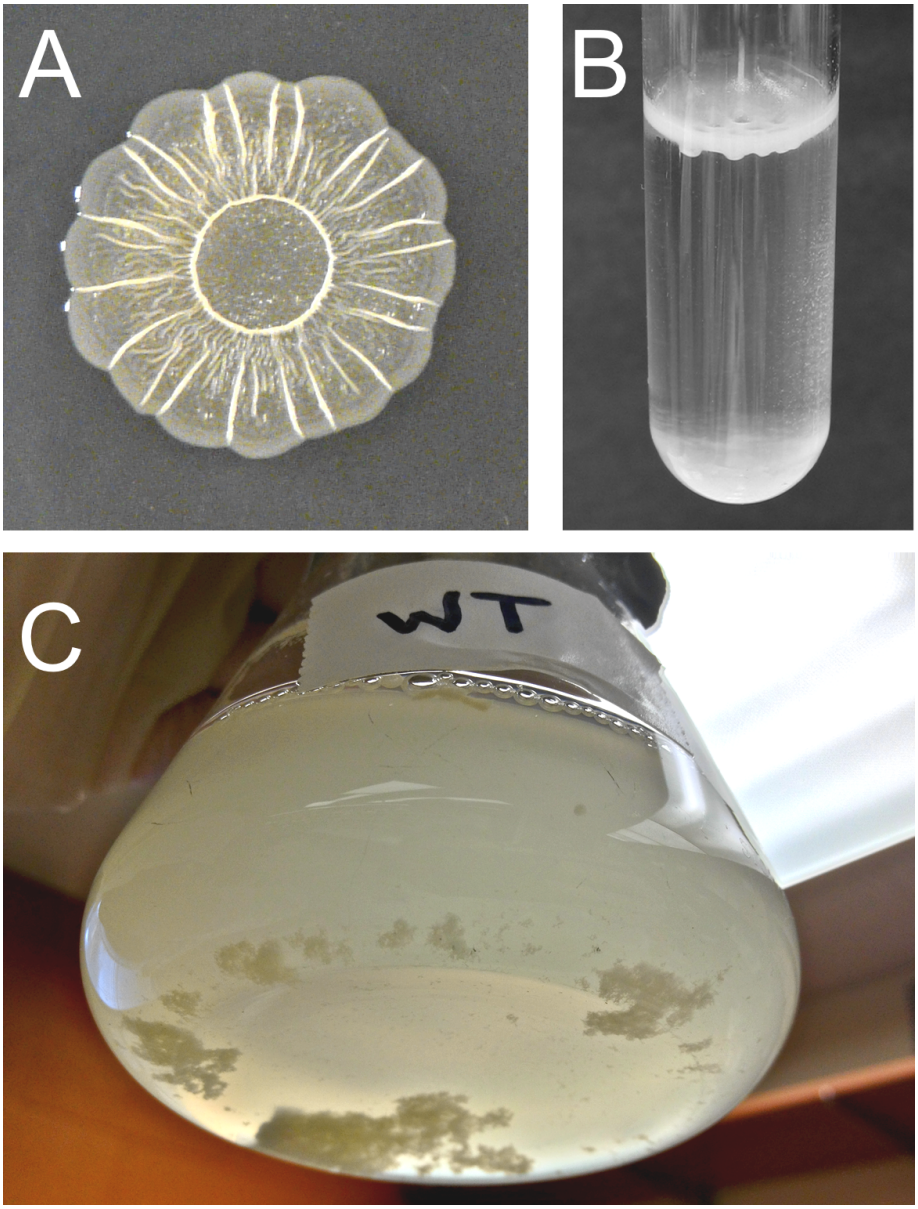

Figure 1.2. Examples of *Salmonella* biofilm formation.

Salmonella biofilm formation has been characterized in the laboratory using several formats of growth. (A) Colonies on 1% tryptone agar form the red, dry, and rough morphotype associated with *Salmonella* biofilm formation. The red appearance of the colony is visible when the media is supplemented with Congo Red dye. (B) Pellicle formation at the air-liquid interface of a liquid culture (adapted from (126)). (C) *Salmonella* form multicellular aggregates and planktonic cells within the bulk liquid phase of a flask culture.

recalcitrant nature of this new fimbrial type through observations that the polymer could only be depolymerized by treatment with 90% formic acid, unlike other fimbriae that were soluble by other means of denaturation (74). Thin aggregative fimbriae in *S.* Enteritidis were originally considered distinct from curli fimbriae in *E. coli* due in part to differences in amino acid residues but mainly to their constitutive expression in *S.* Enteritidis at 37˚C and 28˚C (74). However, this analysis was based on a false identification of the curli subunit, where Olsen *et al.* had in fact identified Crl, a biofilm regulatory protein (named Crl for curli) (87). Upon identifying the genes responsible for the subunits of curli fimbriae, sequence homology revealed significant identity between the genes for thin aggregative fimbriae in *Salmonella* and were curli fimbriae in *E. coli*, hence the adoption of the gene symbol for curli synthesis genes (*csg)* in *Salmonella* (88-90).

At the time of their discovery, curli fimbriae were unique/distinct from other identified pili or fimbriae in that their expression required not only low osmolarity, but also starvation due to extended growth periods and an ambient temperature of less than 32˚C (87). These specific activating conditions suggested the potential for multiple regulatory inputs into expression of curli fimbriae. Using transposon mutagenesis. Hammar *et al.* identified the presence of two polycistronic *csg* operons responsible for curli biosynthesis in *E. coli*, including *csgD*, the main transcriptional regulator demonstrated as necessary for curli biosynthesis (91). The requirement for *csgD* expression for curli fimbriae biosynthesis was later confirmed in *Salmonella* (88, 89). Since then, multiple other transcriptional regulators have been identified that influence biofilm formation through regulation of *csgD* expression, as opposed to direct regulation of the genes responsible for biosynthesis of the biofilm extracellular matrix. The *csgD* promoter has significant sequence homology with other promoters activated by RpoD, the RNA polymerase sigma factor responsible for vegetative cell growth (75). However, limitation of *csgD* expression to stationary phase growth is likely the consequence the *csgD* promoter region being silenced by the histone-like nucleoid structural (H-NS) protein (92-94). Activation of *csgD* expression requires the presence of RpoS, the primary sigma factor during stationary phase growth and regulator of the general stress response. Crl, a DNA-binding transcriptional regulator, facilitates transcription of RpoS-

dependent genes like *csgD* by facilitating efficient open complex formation specifically at lower temperatures (95). In addition, Gerstel *et al.* provided experiment evidence proving regulation of *csgD* expression by OmpR, the response regulator of the EnvZ/OmpR two-component signal transduction system, in response to changes in osmolarity (93). Using DNA footprinting assays, the authors revealed multiple binding sites for OmpR in the *csg* intergenic region. Under conditions of low osmolarity, low cellular concentrations of phosphorylated OmpR bind a high-affinity site directly upstream of the -35 region of the *csgD* promoter and promote expression of biofilm formation. In contrast, conditions of high osmolarity result in high intracellular concentrations of phosphorylated OmpR, which bind low-affinity sites further upstream of the *csgD* promoter and inhibit its expression. CpxAR, an additional two-component signal transduction system, has also been demonstrated to respond to changes in osmolarity. This system has been mainly associated with relaying envelope stress sensed by CpxA to the response regulator CpxR through transfer of a phosphate group. In *E.coli*, CpxR~P has been shown to bind the promoters for both *csg* operons, through its binding sites that either overlap the activating binding site for OmpR in the *csgD* promoter, or exist immediately downstream of the transcriptional start site for *csgBAC* (96). The environmental signals that were initially identified for the expression of curli fimbriae hinges on the involvement of these pivotal transcription factors that regulate *csgD* expression (87).

Several additional regulatory components have since been identified for their effects on *csgD* expression (reviewed in (97)). However, only a few of these factors directly interact with the intergenic region between the *csg* operons, in addition to those factors already described. Integration host factor (IHF) is a histone-like DNA-binding protein that induces DNA topology changes to either activate or repress the expression of the corresponding promoter region. A total of three IHF binding sites have been identified for the *csg* region; IHF1, which when bound by IHF activates *csgD* expression, IHF3, which exists within the *csgD* coding sequence and represses *csgD* transcription when bound by IHF, and IHF2, which does not affect *csgD* transcription when bound (93, 98). An additional transcriptional regulator, MlrA, is also hypothesized to induce DNA topology changes through its binding to a site near IHF1 (99). While no

information exists on factors that influence MlrA activity, a cosmid screen by Brown *et al.* identified this regulator for its specificity on promoting transcriptional activation of *csgDEFG* (100). In addition to this transcriptional regulation, *csgD* also contends with multiple small RNA molecules that inhibit its post-transcriptional expression. Through their interaction with Hfq, an RNA-binding protein, these sRNA target the 5' untranslated region of *csgD* mRNA molecules and can activate or repress translation (101-105). While many of these sRNA molecules are expressed in response to the known environmental conditions that also activate transcription factors (i.e., increased cell envelope stress, high osmolarity, and carbon limitation), some sRNA molecules point to potentially novel biofilm regulation (e.g. GcvB, which is induced by the glycine cleavage system in response to high amino acid supply) (Reviewed (106, 107)). Based on their activating signals and regulatory impact, post-transcriptional modulation by sRNA molecules likely acts as a potent inhibitory cue in the regulation of CsgD synthesis, functioning as a novel and potent response that shuts down biofilm expression once cells sense the presence of an established extracellular matrix structure and a mature state of biofilm development. Altogether, the accumulation of multiple regulatory signals into *csgD* expression and synthesis provides strong evidence that CsgD is the main transcriptional regulator controlling the biosynthesis of extracellular matrix components and biofilm development in *Salmonella*.

Cellulose is recognized as the second major component of the biofilm extracellular matrix. This exopolysaccharide is fundamental in all lab-modeled formats of *Salmonella* biofilm formation (reviewed in (97)). Cellulose facilitates long-distance interactions between *Salmonella* cells and physical adherence of the biofilm to different surfaces (108). In *E. coli* and *S. enterica*, the cellulose biosynthesis machinery is encoded by two divergent operons, *bcsEFG* and *bcsRQABZC* (76, 79, 109) (110). BcsA and BcsB are the main catalytic subunits of the core enzyme complex for cellulose biosynthesis, while the remaining proteins enhance enzyme activity and product yield (110). Cellulose biosynthesis is regulated on both a transcriptional and post-transcriptional level. The *bcs* operons are expressed during biofilm formation, but the BcsAB enzyme complex exists in an autoinhibitory state. Allosteric binding of the PilZ domain of BcsA by the bacterial secondary messenger bis-(3'-5')-cyclic dimeric guanosine monophosphate (c-di-GMP)

activates the enzyme complex, allowing the main substrate, UDP-glucose, to interact with the BcsA active site (110-112).

While c-di-GMP was originally described for its importance in cellulose biosynthesis (112), this secondary messenger molecule is now appreciated for its dynamic role in regulating multiple cell processes (reviewed in (113)). C-di-GMP is predominantly associated with regulating the lifestyle transition between motility and sessility (113). In general, low intracellular levels of c-di-GMP promotes virulence and cell motility, while high levels of c-di-GMP reduce virulence and motility and activate cellulose biosynthesis (97, 113). Diguanylate cyclase enzymes contain an active GGDEF catalytic domain that generates c-di-GMP from two molecules of guanosine triphosphate (GTP). In contrast, phosphodiesterase enzymes contain active EAL or HD-GYP domains that degrade c-di-GMP into 5'-phophoguanyl-(3'-5')-guanosine (pGpG), which is subsequently degraded further into two molecules of guanosine monophosphate (GMP). AdrA was the first diguanylate cyclase associated with cellulose biosynthesis in *Salmonella*, and was identified based on its transcriptional activation by CsgD (108, 109). Since then, a total of 19 proteins has been identified with GGDEF and/or EAL domains (reviewed in (97)). Systematic analysis of these enzymes revealed that only one other diguanylate cyclase, STM1987, can activate cellulose biosynthesis independently of AdrA activity (114). Further, CsgD does not transcriptionally regulate the expression of STM1987 (76, 115, 116). C-di-GMP has also been implicated in regulating CsgD at both the transcriptional and post-transcriptional level. This level of regulation involves a separate set of diguanylate cyclases (i.e., STM2123, STM3388) (117) and phosphodiesterases (i.e., STM1703, STM1827), although some phosphodiesterases are involved in modulating the c-di-GMP pools that regulate either CsgD or cellulose biosynthesis (i.e., STM3611, STM4264) (114). To add to this complexity, the protein STM4551 contains a GGDEF domain and regulates *csgD* transcription, but does not contribute to c-di-GMP synthesis (116). The mechanisms that allow for such hierarchy in c-di-GMP signaling are still elusive, but likely involve several variables, including differential transcription of c-di-GMP metabolizing enzymes, enzyme-specific modulation by different environmental or intracellular signals, co-localization of the

enzymes with their targets (e.g., CsgD, cellulose synthase complex, or the flagellar rotor), and target-specific differences in binding affinity for c-di-GMP (113).

In addition to curli fimbriae and cellulose, CsgD has been directly linked to the transcriptional regulation of two other extracellular matrix components, BapA and the O-antigen capsule. BapA is a large outer membrane surface-associated protein (approximately 386 kilodaltons (kDa)) with homology to a related protein involved in biofilm formation in the gram-positive bacteria *Staphylococcus aureus* (118, 119). *S.* Enteritidis *bapA* mutants form a weak pellicle at the air-liquid interface of broth cultures compared to wildtype strains, suggesting its potential role in *Salmonella* biofilm formation. Overexpression of curli fimbriae allowed for specific refortification of the pellicle structure in this mutant. These results led the authors to hypothesize that BapA strengthens cell-cell interactions within the biofilm, either through direct association of BapA molecules between the cells, or by strengthening curli fimbriae-mediated cellular associations. However, BapA has been demonstrated as nonessential for the rdar morphotype of *Salmonella* grown on agar plates (120). Therefore, further evidence is required to understand the role of BapA in the CsgD regulon and *Salmonella* biofilm formation. In contrast, the O-antigen capsule was identified as an additional unknown exopolysaccharide produced by *Salmonella* during rdar morphotype development on agar plates (78, 121). The O-antigen capsule is composed of 2300 repeating oligosaccharide units that are covalently linked to lipids anchored in the outer membrane (78, 122). CsgD was shown to indirectly activate the expression of genes for the O-antigen capsule by inhibiting the expression of a transcriptional repressor, *yihW* (78). While the O-antigen capsule is not directly required for the cell interactions and the rdar morphotype, it has been proven essential for the survival of *Salmonella* cells exposed to desiccation (78, 80) and has been shown as important for the *Salmonella* attachment to and biofilm formation on alfalfa plants (123). Several other structural components have been suggested as important in certain models of *Salmonella* biofilm formation but have not been linked to regulation by CsgD, and as such fall outside the scope of this discussion (reviewed in (97)).

1.2.2 Genetic and Phenotypic Conservation

Several studies have used the rdar morphotype to evaluate the biofilm phenotype of *Salmonella* serovars obtained from clinical samples, animals, or from meat and produce. Conservation of the biofilm phenotype was highest in *S.* Typhimurium and *S.* Enteritidis isolates associated with gastroenteritis, with 70 to 90% of isolates exhibiting the rdar morphotype (76, 80, 124, 125). When studies were expanded to include all strains of *Salmonella enterica* subspecies *enterica*, isolates that were negative for a biofilm phenotype were associated with invasive diseases in different species, including humans (*S.* Typhi and *S.* Paratyphi A), pigs (*S. Choleraesuis*), chickens (*S.* Gallinarum, and pigeons (*S.* Typhimurium var. Copenhagen). To determine if biofilm formation is present across *S. enterica* and *S. bongori*, White *et al*. performed a screen on *Salmonella* Reference Collection C (126). Their study revealed conservation of the biofilm phenotype in 90% of the 80 isolates representing all 7 *Salmonella* groups. However, *S. enterica* subspecies *arizonae* was the only group where all representative isolates demonstrated a biofilm-negative phenotype. Isolates in this group are frequently isolated from the gut of reptiles and snakes, and therefore may be part of the commensal microflora. Conservation of the biofilm phenotype in invasive nontyphoidal *Salmonella* strains has yet to be determined by large-scale screening. However, preliminary studies of strains belonging to this sequence type have shown that the rdar morphotype is impaired or lost, providing potentially further evidence to the link between gastroenteritis and biofilm formation in *Salmonella* (127). Altogether, assessments of the biofilm phenotype across the genus reveal that biofilm formation is highly conserved in *Salmonella* associated with gastroenteritis, but is lost in *Salmonella* strains that are responsible for invasive disease or are adapted to life in the host.

Curli fimbriae are the only fimbrial type whose operons have been detected in almost all *Salmonella* isolates, suggesting that they serve an important general function in the lifestyle of *Salmonella* (126, 128, 129). Curli operons with nearly identical sequence similarity have also been found in other bacteria belonging to the family *Enterobacteriaceae*, including *E. coli*, *Citrobacter, Enterobacter,* and *Klebsiella* (89, 130, 131). While the *csg* operons demonstrate significant sequence conservation across different enterobacteria, White *et al.* found that the intergenic region between the *csg*

operons has an overall sequence identity of 67% across *Salmonella* groups, suggesting significant diversity in the promoter region for either *csg* operon and potential differences in biofilm regulation across the genus (126). Indeed, the majority of mutations within the *csg* intergenic regions for the *Salmonella* isolates were localized to areas that are not adjacent to the *csg* promoters. Transcriptional reporter fusions for the *csgD* and *csgB* promoters of *Salmonella* isolates from all 7 subgroups revealed conserved transcriptional expression for both operons. Further, sequence analysis of the *csg* promoter regions revealed that the -10 and -35 regions of both *csg* promoters were almost entirely conserved across all *Salmonella* isolates. In comparing the *csg* intergenic regions between *E. coli* and *Salmonella*, Gerstel and colleagues found a similar pattern of sequence conservation, where sequence identity between the species was greatest in the promoter regions immediately upstream for either operon that contain pivotal regulatory binding sites (93).

In the studies reviewing phenotypic conservation of biofilm formation, some of the *Salmonella* isolates were found to have *trans* differences corresponding to altered regulation of the *csg* operons (124, 126). Many of these strains were found to have mutations affecting the *rpoS* allele, however, other strains were suspected of having mutations in other pathways involved in the regulation of *csgD* expression. White *et al.* found over-representation of such *trans* mutations in a small sub-collection of isolates representing the diversity of strains in the seven *Salmonella* groups (126). The authors postulated that the over-representation of *trans* mutations was most likely due to domestication of these isolates, a phenomenon that has been demonstrated in repeated sub-culturing of several bacteria including *Salmonella*, *Mycobacterium bovis* and *E. coli* (132-135). Subsequent work by Davidson *et al.* established that loss of rdar biofilm formation was common during laboratory passage. In their studies, they demonstrated that passaging *S.* Typhimurium cultures in stationary phase resulted in an increased mutation rate in the *rpoS* allele, while passaging during exponential phase produced other mutations resulting in loss of the rdar morphotype. While expression of the RpoS regulon aids cells by protecting against multiple stresses, expression of the RpoD regulon aids cells with nutrient scavenging and likely provides a competitive edge during laboratory-based growth.

The notion that expression of the *csg* operons is conserved across the Salmonellae is contradicted by rare exceptions. For *S. arizonae* isolates, *cis* mutations in the CsgD binding site for the *csgB* promoter and in conserved nucleotides for OmpR binding site required for *csgD* expression corresponded with loss of transcriptional activity and the biofilm phenotype (126). In *S.* Typhimurium var. Copenhagen isolates, the prevalent mutation was a G to T transversion in the -35 region, which was demonstrated to significantly reduce *csgD* promoter activity and may partially explain the loss of the biofilm phenotype in these invasive Salmonellae (124). For *S.* Typhi isolates, the intergenic region has conserved sequence, but multiple mutations exist within the *csg* and *bcs* operons, where preliminary stop codons within *csgD* and *bcsC* genes may eliminate the possibility for synthesis of curli fimbriae and cellulose altogether (124, 136). Similarly, a few mutations have been shown to correspond with increased biofilm biogenesis in *Salmonella*. For *S.* Enteritidis 3b and *S.* Typhimurium SR-11b, a specific G to T transversion in the activating OmpR site of the *csgD* promoter allows for *csgD* expression by RpoD instead of RpoS, resulting in a constitutive biofilm phenotype that is expressed in a growth- and temperature-independent manner (75, 77). White *et al.* hypothesized that *cis* mutations in the *csg* promoters are more likely to indicate an evolution in the lifestyle of *Salmonella* isolates (126).

1.2.3 The CsgD Regulon

The integration of the CsgD regulon with other global regulatory elements shows its potential to exert profound changes in cellular physiology. To quantify the impact of CsgD-regulated extracellular matrix production on cell physiology, White *et al.* compared the metabolic and transcriptional profiles of wildtype and *csgD* mutant *S.* Typhimurium cells (137). The metabolites extracted from wildtype, biofilm-forming cells were associated with the end products of gluconeogenesis, osmoprotectants, and other products related to defenses against reactive oxygen species. In contrast, *csgD* mutant cells were inhibited for gluconeogenesis, as evident through elevated levels of adenine monophosphate (AMP), polyamines, and molecules of the upper tricarboxylic acid (TCA) cycle. These metabolites indicated novel physiology for wildtype biofilm cells. The presence of osmoprotectants indicated signals of high osmolarity for biofilm cells,

which was unexpected because the cells were grown in the low osmolarity conditions of 1% tryptone media (89, 93, 137). These findings together indicated that the metabolic requirements for biofilm production and the presence of the mature extracellular matrix may indirectly influence the global gene expression within a cell (137). To determine if cell aggregation had a global effect on gene expression at the transcriptional level, White *et al.* generated a library of transcriptional fusion constructs for genes in pathways that synthesize the metabolic products identified in wildtype and *csgD* mutant cells. Most of the genes included in their reporter library were expressed in a similar temporal pattern, with their maximal expression beginning in sync with high activity of RpoS sigma factor and the expression of *csgDEFG*, *csgBAC*, and *adrA* promoters. In contrast, expression of the reporter library in *csgD* mutant cells had lost this ordered temporal expression. Altogether, these observations suggested that cells within multicellular aggregates might have gene expression occurring in anticipation of extracellular matrix production or responding to changes in the local microenvironment created by the presence of the biofilm superstructure (137).

1.2.4 CsgD Bistability and Phenotypic Variation

Biofilm development represents an important driver of phenotypic heterogeneity amongst otherwise genetically identical cells (179). While some diversity arises from the unique niches created within the biofilm's complex physical structure, probes into the genetic processes governing biofilm formation have revealed the potential for population heterogeneity long before synthesis of the extracellular matrix. Indeed, the complex interplay of environmental signaling and global transcriptional regulators creates the potential for cell-to-cell variation in the expression of *csgD*. White *et al.* (2008) first observed the presence of two populations of cells within *S.* Typhimurium liquid cultures grown under biofilm-inducing conditions (Fig 1.2c). One group of cells were aggregated and had high curli expression, whereas the other group of cells remained as single cells and had low curli expression. Grantcharova *et al.* examined this phenomenon at the molecular level (138) and were able to monitor dynamic changes in CsgD synthesis by generating a translational fusion of CsgD to green fluorescent protein (GFP^+) and tracking using fluorescence microscopy and fluorescence-assisted cell sorting (FACS).

The authors observed a state of genetic bistability during their experiments, where unique subpopulations of cells with high CsgD synthesis or where CsgD synthesis was absent were consistently identified across three models of biofilm growth (spot cultures on agar plates, in static liquid cultures, or during growth in a flow cell). Cell sorting revealed that at maximum, 52% of cells in the total population existed in a "CsgD-ON" state after 24 hours of growth in liquid media. The CsgD-ON cells were demonstrated to be the main sources of cellulose production, and were associated with other CsgD-synthesizing cells as multicellular aggregates that had sedimented out of the culture medium or had attached to an abiotic glass surface. In contrast, cells in a "CsgD-OFF" state existed as singular, motile planktonic cells within the bulk liquid phase of the cultures and were not associated with multicellular aggregates. Similarly, mutant cells lacking the *csgD* gene also existed as single cells and had 40% less cell attachment to the glass surfaces of a flask or flow cell compared to CsgD-ON wildtype cells. The authors concluded that this defect in cellulose production and attachment in the *csgD* mutant was due to a lack of activation of *adrA* expression, which normally contributes a significant amount to the c-di-GMP pool that promotes cellulose biosynthesis. Overall, their findings suggest that *csgD* expression is an all-or-none event, and that variation in its expression at the single cell level generates disparate cell phenotypes with unique physiology in an otherwise isogenic population of nontyphoidal *Salmonella* cells.

One of the requirements for bistable gene expression is the presence of an auto-regulatory loop involving the gene in question, pushing its expression past intermediate levels and towards an extreme "ON" state (139). While evidence for such regulation has yet to be proven in *Salmonella*, a potential auto-regulatory loop has been shown in a study observing biofilm formation in *E. coli* (140). In this study, Gualdi *et al.* used variants of the biofilm-negative *E. coli* strain MG1655 that were complemented for the biofilm phenotype in either a growth-phase dependent manner via a mutated *ompR* allele, or through constitutive, low-level *csgD* expression from a plasmid. After 18 hours of growth at 30˚C, a comparison of the proteomes of these strains with respect to the wildtype parental strain identified thec differential expression of several proteins encoded by genes belonging regulated by the stationary phase sigma factor RpoS. Constitutive expression of *csgD* was unable to change the synthesis of these proteins without the

presence of a functional *rpoS* allele. Similarly, the combination of *csgD* and *rpoS* was required to promote significant up-regulated transcription of an examined selection of genes from the *rpoS* regulon. Therefore, the authors reasoned that CsgD is able to affect the expression of *rpoS*-dependent genes by modulating the concentration or activity of the RpoS sigma factor. Previous work by the same research group had shown CsgD-dependent up-regulation of *iraP*, which encodes a protein that was later demonstrated to stabilize the protein levels of RpoS sigma factor by inhibiting its degradation by the RssB-ClpXP protein complex (141, 142). Gualdi and co-workers were able to provide direct evidence of the link between CsgD, IraP, and RpoS sigma factor by demonstrating the reduced half-life of RpoS sigma factor in an *iraP* mutant with constitutive CsgD synthesis. Further, deletion of the *iraP* allele from the constitutive *csgD* strain abrogated the changes in protein synthesis the authors had previously shown to be connected to expression of *csgD*. Altogether, these results suggest that CsgD affects the *rpoS* regulon, of which it is a member, by indirect stabilization of the intracellular levels of RpoS sigma factor via *iraP* activation.

1.3 Tn7 Transposition and Biotechnology

1.3.1 Tn7 Transposition Biology

Transposons are mobile genetic elements that move from one position to another in the genome of a cell (143). These systems are broadly classified into two categories: retrotransposons, which rely on reverse transcription to move between segments of DNA via an RNA intermediate, and DNA transposons, which are mobilized directly from DNA to DNA (144). Tn7 is a bacterial DNA transposon that is unique due to the site-specific nature of its integration into the bacterial chromosome (145). This transposon is comprised of the genes responsible for its transposition, *tnsABCDE*, a variable region containing additional genes, and two bordering sequences at either end of the element, Tn7L and Tn7R (146). Tn7 was first described for its role in conferring trimethoprim resistance to *E. coli* strains associated with calves (145, 147). However, increasing availability of bacterial genome sequences has revealed that Tn7 and Tn7-related

elements are ubiquitous throughout the bacterial kingdom, with significant gene diversity represented in the transposon's variable region (148).

One potential reason for the widespread adoption of Tn7 transposition systems across bacteria is the ability to target a conserved chromosomal sequence and integrate into the genome without affecting the host's gene repertoire (149, 150). TnsD is the target specificity factor that directs site-specific integration of Tn7 into the chromosome (149). Research into the mechanism of Tn7 integration has shown that TnsD is a DNA-binding protein that recognizes a conserved sequence found in the 3' end of the highly conserved *glmS* gene (151). Two host proteins, the acyl carrier protein (ACP) and L29 (a nonessential component of the ribosomal large subunit), stimulate binding of TnsD to this region (152). Interaction of TnsD with the target sequence induces a conformational change in the target DNA, which recruits an additional DNA-binding protein, TnsC (153). The formation of this nucleoprotein structure promotes integration of the Tn7 element into the *att*Tn7 site, located 25 basepairs downstream of the TnsD binding region (154). This site of integration exists within the transcriptional terminator sequence for the *glmS* gene (150). However, Tn7 insertion does not affect the function of the encoded enzyme, glucosamine-6-phosphate, which has an essential role in the biosynthesis of the bacterial cell wall (150). Tn7 elements have been demonstrated to integrate into the chromosome at a high frequency, with approximately 10^{-1} and 10^{-2} chromosomal insertions for each donor DNA plasmid (155).

In addition to its role in targeting Tn7 transposition to the *attTn7* site, TnsC also regulates the activity of the Tn7 transposase machinery involved in transposon excision, TnsAB. Once the nucleoprotein complex (i.e., TnsABCD) is established downstream of *glmS*, Tn7 transposition into the chromosome occurs via a "cut and paste" mechanism that involves the activities of both TnsA and TnsB (156). TnsA shares structural similarity to a restriction endonuclease (157), and is responsible for nicking the 5' end of both strands of the donor DNA molecule (156, 158, 159). In contrast, the TnsB enzyme is a member of the retroviral integrase superfamily, and is responsible for breaking the 3' hydroxyl bonds on both strands of donor DNA (159). The cleavage of both DNA strands results in complete liberation of the transposon element from the donor DNA, leaving

behind a double-stranded DNA break that is repaired by the host's machinery via homologous recombination (156).

TnsA and TnsB also facilitate integration of the transposon element into the target DNA sequence. TnsB recognizes specific binding sites found within the Tn7L and Tn7R regions on either end of the transposon (160, 161). It is hypothesized that differences in both the number and configuration of TnsB binding sites between Tn7L and Tn7R ensures insertion of the Tn7 element into the target DNA in a specific "left-to-right" orientation (reviewed in (162)). The transposon ends are then joined together by TnsA, forming the paired end DNA complex that is necessary for transposition (163). Following this complex formation, TnsB facilitates "strand transfer", or the integration of the 3'-hydroxyl ends of the Tn7 element into the target DNA sequence (156, 159). These end sequences are integrated 5 basepairs apart in the target DNA sequences, creating a single-stranded gap on either end of the integrated transposon element (156). Reparation of these gaps by the host machinery results in 5-basepair duplications that flank either end of the transposon sequence (154, 156).

While TnsD-mediated Tn7 transposition facilitates vertical transfer of the transposon to bacterial progeny, a second targeting factor, TnsE, is important for horizontal dissemination of Tn7 elements across bacterial species. Unlike TnsD, TnsE-mediated transposition relies on identifying a DNA structure instead of a specific DNA sequence (164). TnsE associates with sites of lagging strand synthesis during DNA replication, and will simultaneously bind DNA structures with a 3' recessed end and the sliding processivity factor, DnaN (164, 165). This unique feature of TnsE targeting explains why TnsE-mediated transposition frequently results in integration of Tn7 elements into conjugal plasmids (166). Once mobilized, these plasmids enter the new cell as single-stranded DNA, and are replicated by the host DNA replication machinery in a manner similar to lagging strand synthesis (167). TnsE recognizes the DNA structure associated with this synthesis and will incorporate the Tn7 element into the plasmid DNA sequence at random positions (166). Due to recognition of a specific DNA structure, TnsE can also facilitate Tn7 incorporation into the M13 bacteriophage or the bacterial chromosome (149, 180). However, integration into the chromosome occurs at a much lower frequency than integration into conjugal plasmids (162).

1.3.2 Tn7 Biotechnology

Chromosomal integration of a gene of interest overcomes many of the current pitfalls associated with plasmid-based gene expression. For example, single-copy chromosomal insertion of the gene of interest addresses the issues of plasmid copy number and gene dosage effects that can result in non-physiological levels of gene expression (150, 168). Knodler *et al.* demonstrated that introduction of a cloning vector can also result in unintended secondary effects on other bacterial cell processes, such as *Salmonella* virulence (169). Stable chromosomal integration of the gene eliminates the issue of plasmid loss associated with imperfect plasmid segregation (150). In some cases, the selection pressure required for plasmid maintenance may even be impossible or undesirable, such as during experiments with bacterial biofilm formation or *in vivo* studies of bacteria in animals or humans (168, 170, 171). Therefore, Tn7 transposition is an attractive option for generating a broadly applicable molecular tool that would facilitate site-specific gene insertion into the bacterial chromosome.

One of the earliest and most widely-adopted Tn7 biotechnologies was developed by Bao and colleagues in 1991 (172). Their plasmid-based system was designed to establish a stable chromosomal integrant, where the gene of interest would be permanently incorporated into the targeted chromosomal region downstream of *glmS*. To achieve this, Bao *et al.* devised a two-plasmid system involving a "helper" plasmid (pUX-BF13) carrying the Tn7 transposase genes, and a "carrier" plasmid (pUX-BF5) with a miniaturized Tn7 element. This mini-Tn7 element consisted of a kanamycin resistant cassette flanked by the Tn7L and Tn7R end sequences; researchers could use one of the available restriction sites to insert their gene of interest into this transposon. Both the carrier and helper plasmids were then simultaneously conjugated or electroporated into the targeted bacterial strain. One of the key features critical to this molecular system was to develop a "suicide vector" strategy, where both plasmids would be naturally lost following successful chromosomal integration of the gene of interest. To limit maintenance of either plasmid in the targeted bacterial strain, Bao *et al.* cloned the transposase genes and mini-Tn7 element into vectors associated with a narrow bacterial host range. The *tnsABCDE* genes were ligated to a plasmid with an R6K origin, which can only be replicated in host strains that carry the lambda *pir* gene. In contrast, the

carrier plasmid was based on a ColE1 vector that can only replicate in Enterobacterial species such as *E. coli* and *Salmonella*.

The Tn7 transposition system developed by Bao *et al.* was a significant advancement over the random chromosomal insertion strategy provided by existing Tn5-based technologies. However, the use of this system was hampered by three important factors: (1) a lack of unique restriction sites available for cloning into the mini-Tn7 element (173), (2) the potential for random chromosomal insertion due to the *tnsE* gene (174), and (3) the inability to use this suicide-based approach in Enterobacteriaceae due to the ColE1 vector backbone (168, 170, 175). Choi and colleagues addressed these limitations by introducing several modifications to this two-plasmid system (170). To ensure site-specific integration of the mini-Tn7 element, Choi *et al.* deleted a large portion of the *tnsE* 5' coding sequence from pUX-BF13, resulting in a new helper plasmid, pTNS1. Many additional modifications affected the functionality of the carrier plasmid and its mini-Tn7 element. Choi *et al.* made it possible to use this Tn7 system in Enterobacteriaceae by replacing the ColE1 origin of replication with an R6K origin. Addition major modifications were made to the mini-Tn7 element itself. First, the researchers inserted T_0 and T_1 transcriptional terminators to the Tn7R end of the element to act as a replacement for the native *glmS* transcriptional terminator that is interrupted by Tn7 integration. Second, Choi *et al.* included an expansive multiple cloning site to facilitate insertion of genes of interest into the mini-Tn7 element. To further broaden the application of this system, Choi *et al.* also generated a family of pre-built mini-Tn7 elements. These elements provided researchers with the option to select a desired antibiotic resistance marker, generate transcriptional fusions with *lacZ* or bacterial luciferase genes, or insert a gene into a multiple cloning site downstream of an IPTG-inducible *tac* promoter. Unlike previous Tn7 systems (172, 173, 176), their mini-Tn7 element also included FRT sites flanking the antibiotic resistance cassettes to allowing for excision of this cassette by yeast flippase (FLP) following chromosomal integration of the transposon and verification of selected bacterial clones. Altogether, the modifications by Choi *et al*. vastly improved the robustness and functionality of the original Tn7 system.

Following the release of this toolset, subsequent iterations of the Tn7 system focused on improving the efficiency of successful transposition events. To increase expression of the transposition genes, Choi *et al.* introduced additional helper plasmids where the native promoter for *tnsABCD* genes had been replaced by the constitutively-expressed *lac* promoter (pTNS2) (170) or the P1 integron and *lac* promoters in tandem (pTNS3) (177). Despite the system's extensive use in some bacterial species (i.e., *Pseudomonas* aeruginosa), there were only rare reports of its use in *Salmonella* and *E. coli*. We hypothesized that the suicide vector approach of Choi *et al*'s system may not allow for sufficient expression of the *tnsABCD* genes, resulting in failure to integrate the mini-Tn7 element prior to loss of the helper plasmid. During the course of this thesis, Crépin and colleagues published an alternative approach to increasing transposase expression with a new helper plasmid, pSTNSK (168). Their two-plasmid system involved the transfer of the *tnsABCD* genes from pTNS2 to a temperature-sensitive plasmid (i.e., pSC101 origin), allowing for conditional maintenance of the *tns* genes within a bacterial strain. The authors reported significant success with this system in a range of Enterobacterial strains, observing transposition efficiencies between 68 and 96%. Similarly, researchers McKenzie and Craig (pGRG25) incorporated both aspects of improved promoter strength for the *tnsABCD* genes and conditional maintenance of the Tn7 system in their one-plasmid system, pGRG25 (150). In contrast to the system devised by Crépin *et al.*, this tool positioned both the transposase genes and the mini-Tn7 element together on a temperature-sensitive plasmid. Further, McKenzie and Craig's system allowed the user to modulate expression levels of the transposition genes by placing them under the control of an arabinose-inducible *araBAD* promoter. In experiments with *E. coli* K-12, the authors demonstrated that use of their system resulted in a transposition efficiency of 79% (150). However, both our research group and Crépin *et al.* were unsuccessful with using this single-plasmid system in pathogenic *E. coli* and *Salmonella* (168, 178). These observations suggest that further improvements must be made to the available Tn7 systems in order to generate robust tools that allow for chromosomal integration across multiple bacterial species and strains.

1.4 References

1. **Su L-H**, **Chiu C-H**. 2007. *Salmonella*: clinical importance and evolution of nomenclature. Chang Gung Med J **30**:210–219.

2. **Public Health Agency of Canada**. 2010. *Salmonella enterica* spp. - Pathogen Safety Data Sheets. http://www.phac-aspc.gc.ca/lab-bio/res/psds-ftss/salmonella-ent-eng.php

3. **Crosa JH**, **Brenner DJ**, **Ewing WH**, **Falkow, S.** 1973. Molecular relationships among the Salmonelleae. *Journal of Bacteriology*, *115*(1), pp.307-315.

4. **Nataro JP**, **Bopp CA**, **Fields PI**, **Kaper JB**, **Strockbine NA**. 2007. *Escherichia*, *Shigella*, and *Salmonella*, pp. 671–687. *In* Manual of Clinical Microbiology, 9 ed. Washington: ASM Press.

5. **Tindall BJ**. 2005. Nomenclature and taxonomy of the genus *Salmonella*. International Journal of Systematic and Evolutionary Microbiology **55**:521–524.

6. **Le Minor L**, **Popoff MY**. 1987. Designation of *Salmonella* enterica sp. nov., nom. rev., as the Type and Only Species of the Genus *Salmonella*: Request for an Opinion. International Journal of Systematic and Evolutionary Microbiology **37**:465–468.

7. **Grimont P**, **Weill FX**. 2007. Antigenic formulae of the *Salmonella* serotypes. WHO Collaborating Centre for Reference and Research on *Salmonella*.

8. **Dieckmann R**, **Malorny B**. 2011. Rapid Screening of Epidemiologically Important *Salmonella enterica* subsp. *enterica* Serovars by Whole-Cell Matrix-Assisted Laser Desorption Ionization-Time of Flight Mass Spectrometry. Applied and Environmental Microbiology **77**:4136–4146.

9. **CDC**. 2011. National *Salmonella* Surveillance Overview 1–12.

10. **Waldner L**, **MacKenzie K**, **Köster W**, **White A**. 2012. From Exit to Entry: Long-term Survival and Transmission of *Salmonella*. Pathogens 2012, Vol 1, Pages 128-155 **1**:128–155.

11. **Brenner FW**, **Villar RG**, **Angulo FJ**, **Tauxe R**, **Swaminathan B**. 2000. *Salmonella* nomenclature. Journal of Clinical Microbiology **38**:2465–2467.

12. **Betancor L**, **Yim L**, **Martínez A**, **Fookes M**, **Sasias S**, **Schelotto F**, **Thomson N**, **Maskell D**, **Chabalgoity JA**. 2012. Genomic Comparison of the Closely Related *Salmonella enterica* Serovars Enteritidis and Dublin. Open Microbiol J **6**:5–13.

13. **Kirk MD**, **Pires SM**, **Black RE**, **Caipo M**, **Crump JA**, **Devleesschauwer B**, **Döpfer D**, **Fazil A**, **Fischer-Walker CL**, **Hald T**, **Hall AJ**, **Keddy KH**, **Lake**

RJ, **Lanata CF**, **Torgerson PR**, **Havelaar AH**, **Angulo FJ**. 2015. World Health Organization Estimates of the Global and Regional Disease Burden of 22 Foodborne Bacterial, Protozoal, and Viral Diseases, 2010: A Data Synthesis. PLoS Med **12**:e1001921–21.

14. **Galanis E**, **Fo Wong Lo DMA**, **Patrick ME**, **Binsztein N**, **Cieslik A**, **Chalermchikit T**, **Aidara-Kane A**, **Ellis A**, **Angulo FJ**, **Wegener HC**, **World Health Organization Global Salm-Surv**. 2006. Web-based surveillance and global *Salmonella* distribution, 2000-2002. Emerg Infect Dis **12**:381–388.

15. **Reddy EA**, **Shaw AV**, **Crump JA**. 2010. Community-acquired bloodstream infections in Africa: a systematic review and meta-analysis. The Lancet Infectious Diseases **10**:417–432.

16. **Ao TT**, **Feasey NA**, **Gordon MA**, **Keddy KH**, **Angulo FJ**, **Crump JA**. 2015. Global Burden of Invasive Nontyphoidal *Salmonella* Disease, 2010. Emerg Infect Dis **21**:941–949.

17. **Crump JA**, **Sjölund-Karlsson M**, **Gordon MA**, **Parry CM**. 2015. Epidemiology, Clinical Presentation, Laboratory Diagnosis, Antimicrobial Resistance, and Antimicrobial Management of Invasive *Salmonella* Infections. Clinical Microbiology Reviews **28**:901–937.

18. **Feasey NA**, **Dougan G**, **Kingsley RA**, **Heyderman RS**, **Gordon MA MD**. 2012. Invasive non-typhoidal *Salmonella* disease: an emerging and neglected tropical disease in Africa. The Lancet **379**:2489–2499.

19. **Gordon MA**, **Graham SM**, **Walsh AL**, **Wilson L**, **Phiri A**, **Molyneux E**, **Zijlstra EE**, **Heyderman RS**, **Hart CA**, **Molyneux ME**. 2008. Epidemics of invasive *Salmonella enterica* serovar Enteritidis and *S. enterica* Serovar Typhimurium infection associated with multidrug resistance among adults and children in Malawi. Clinical Infectious Diseases **46**:963–969.

20. **Eng S-K**, **Pusparajah P**, **Ab Mutalib N-S**, **Ser H-L**, **Chan K-G**, **Lee L-H**. 2015. *Salmonella*: A review on pathogenesis, epidemiology and antibiotic resistance. Frontiers in Life Science **8**:284–293.

21. **Ochiai RL**, **Acosta CJ**. 2008. A study of typhoid fever in five Asian countries: disease burden and implications for controls. Bulletin of the World Health Organization **86**:260–268.

22. **Hohmann EL**. 2001. Nontyphoidal salmonellosis. Clinical Infectious Diseases **32**:263–269.

23. **Sánchez-Vargas FM**, **Abu-El-Haija MA**, **Gómez-Duarte OG**. 2011. *Salmonella* infections: An update on epidemiology, management, and prevention. Travel Medicine and Infectious Disease **9**:263–277.

24. **Humphries RM**, **Linscott AJ**. 2015. Laboratory Diagnosis of Bacterial Gastroenteritis. Clinical Microbiology Reviews **28**:3–31.

25. **Parry CM**, **Hien TT**, **Dougan G**, **White NJ**, **Farrar JJ**. 2002. Typhoid fever. N Engl J Med **347**:1770–1782.

26. **Peters RPH**, **Zijlstra EE**, **Schijffelen MJ**, **Walsh AL**, **Joaki G**, **Kumwenda JJ**, **Kublin JG**, **Molyneux ME**, **Lewis DK**. 2004. A prospective study of bloodstream infections as cause of fever in Malawi: clinical predictors and implications for management. Trop Med Int Health **9**:928–934.

27. **Morpeth SC**, **Ramadhani HO**, **Crump JA**. 2009. Invasive Non-Typhi *Salmonella* Disease in Africa. Clinical Infectious Diseases **49**:606–611.

28. **Agbor TA**, **McCormick BA**. 2011. *Salmonella* effectors: important players modulating host cell function during infection. Cellular Microbiology **13**:1858–1869.

29. **Moest TP**, **Méresse S**. 2013. *Salmonella* T3SSs: successful mission of the secret(ion) agents. Current Opinion in Microbiology **16**:38–44.

30. **Wisner A**, **Desin T**, **White A**, **Potter A**, **Köster W**. The *Salmonella* Pathogenicity Island-1 and-2 Encoded Type III Secretion Systems.

31. **Costa TRD**, **Felisberto-Rodrigues C**, **Meir A**, **Prevost MS**, **Redzej A**, **Trokter M**, **Waksman G**. 2015. Secretion systems in Gram-negative bacteria: structural and mechanistic insights. Nature Publishing Group **13**:343–359.

32. **Ochman H**, **Groisman EA**. 1995. The evolution of invasion by enteric bacteria. Can J Microbiol **41**:555–561.

33. **Hensel M**. 2004. Evolution of pathogenicity islands of *Salmonella enterica*. Int J Med Microbiol **294**:95–102.

34. **Thiennimitr P**, **Winter SE**, **Bäumler AJ**. 2012. *Salmonella*, the host and its microbiota. Current Opinion in Microbiology **15**:108–114.

35. **Misselwitz B**, **Barrett N**, **Kreibich S**, **Vonaesch P**, **Andritschke D**, **Rout S**, **Weidner K**, **Sormaz M**, **Songhet P**, **Horvath P**, **Chabria M**, **Vogel V**, **Spori DM**, **Jenny P**, **Hardt W-D**. 2012. Near Surface Swimming of *Salmonella* Typhimurium Explains Target-Site Selection and Cooperative Invasion. PLoS Pathog **8**:e1002810–19.

36. **Lara-Tejero M**, **Galán JE**. 2009. *Salmonella enterica* Serovar Typhimurium Pathogenicity Island 1-Encoded Type III Secretion System Translocases Mediate Intimate Attachment to Nonphagocytic Cells. Infection and Immunity **77**:2635–2642.

37. **Misselwitz B**, **Kreibich SK**, **Rout S**, **Stecher B**, **Periaswamy B**, **Hardt WD**. 2010. *Salmonella enterica* Serovar Typhimurium Binds to HeLa Cells via Fim-Mediated Reversible Adhesion and Irreversible Type Three Secretion System 1-Mediated Docking. Infection and Immunity **79**:330–341.

38. **Altier C**. 2005. Genetic and environmental control of *Salmonella* invasion. J Microbiol **43 Spec No**:85–92.

39. **Izoré T**, **Job V**, **Dessen A**. 2011. Biogenesis, Regulation, and Targeting of the Type III Secretion System. Structure/Folding and Design **19**:603–612.

40. **van der Heijden J**, **Finlay BB**. 2012. Type III effector-mediated processes in *Salmonella* infection. Future Microbiology **7**:685–703.

41. **Srikanth CV**, **Mercado-Lubo R**, **Hallstrom K**, **McCormick BA**. 2011. *Salmonella* effector proteins and host-cell responses. Cellular and Molecular Life Sciences **68**:3687–3697.

42. **Fabrega A**, **Vila J**. 2013. *Salmonella enterica* Serovar Typhimurium Skills To Succeed in the Host: Virulence and Regulation. Clinical Microbiology Reviews **26**:308–341.

43. **Raffatellu M**, **George MD**, **Akiyama Y**, **Hornsby MJ**, **Nuccio S-P**, **Paixao TA**, **Butler BP**, **Chu H**, **Santos RL**, **Berger T**, **Mak TW**, **Tsolis RM**, **Bevins CL**, **Solnick JV**, **Dandekar S**, **Bäumler AJ**. 2009. Lipocalin-2 resistance confers an advantage to *Salmonella enterica* serotype Typhimurium for growth and survival in the inflamed intestine. Cell Host & Microbe **5**:476–486.

44. **Winter SE**, **Thiennimitr P**, **Winter MG**, **Butler BP**, **Huseby DL**, **Crawford RW**, **Russell JM**, **Bevins CL**, **Adams LG**, **Tsolis RM**, **Roth JR**, **Bäumler AJ**. 2010. Gut inflammation provides a respiratory electron acceptor for *Salmonella*. Nature **467**:426–429.

45. **Winter SE**, **Thiennimitr P**, **Winter MG**, **Butler BP**, **Huseby DL**, **Crawford RW**, **Russell JM**, **Bevins CL**, **Adams LG**, **Tsolis RM**, **Roth JR**, **Bäumler AJ**. 2010. Gut inflammation provides a respiratory electron acceptor for *Salmonella*. Nature **467**:426–429.

46. **Thiennimitr P**, **Winter SE**, **Winter MG**, **Xavier MN**, **Tolstikov V**, **Huseby DL**, **Sterzenbach T**, **Tsolis RM**, **Roth JR**, **Bäumler AJ**. 2011. Intestinal inflammation allows *Salmonella* to use ethanolamine to compete with the microbiota. Proceedings of the National Academy of Sciences of the United States of America **108**:17480–17485.

47. **McClelland M**, **Sanderson KE**, **Clifton SW**, **Latreille P**, **Porwollik S**, **Sabo A**, **Meyer R**, **Bieri T**, **Ozersky P**, **McLellan M**, **Harkins CR**, **Wang C**, **Nguyen C**, **Berghoff A**, **Elliott G**, **Kohlberg S**, **Strong C**, **Du F**, **Carter J**, **Kremizki C**, **Layman D**, **Leonard S**, **Sun H**, **Fulton L**, **Nash W**, **Miner T**, **Minx P**,

Delehaunty K, **Fronick C**, **Magrini V**, **Nhan M**, **Warren W**, **Florea L**, **Spieth J**, **Wilson RK**. 2004. Comparison of genome degradation in Paratyphi A and Typhi, human-restricted serovars of *Salmonella enterica* that cause typhoid. Nature Genetics **36**:1268–1274.

48. **Dougan G**, **Baker S**. 2014. *Salmonella enterica* Serovar Typhi and the Pathogenesis of Typhoid Fever. Annu Rev Microbiol **68**:317–336.

49. **Gal-Mor O**, **Boyle EC**, **Grassl GA**. 2014. Same species, different diseases: how and why typhoidal and non-typhoidal *Salmonella enterica* serovars differ. Front Microbiol **5**:391.

50. **Gunn JS**, **Marshall JM**, **Baker S**, **Dongol S**, **Charles RC**, **Ryan ET**. 2014. *Salmonella* chronic carriage: epidemiology, diagnosis, and gallbladder persistence. Trends Microbiol **22**:648–655.

51. **Gordon MA**. 2008. *Salmonella* infections in immunocompromised adults. Journal of Infection **56**:413–422.

52. **Winter SE**, **Raffatellu M**, **Wilson RP**, **Russmann H**, **Bäumler AJ**. 2008. The *Salmonella enterica* serotype Typhi regulator TviA reduces interleukin-8 production in intestinal epithelial cells by repressing flagellin secretion. Cellular Microbiology **10**:247–261.

53. **Crawford RW**, **Wangdi T**, **Spees AM**, **Xavier MN**, **Tsolis RM**, **Bäumler AJ**. 2013. Loss of very-long O-antigen chains optimizes capsule-mediated immune evasion by *Salmonella enterica* serovar Typhi. mBio **4**.

54. **Wangdi T**, **Lee C-Y**, **Spees AM**, **Yu C**, **Kingsbury DD**, **Winter SE**, **Hastey CJ**, **Wilson RP**, **Heinrich V**, **Bäumler AJ**. 2014. The Vi Capsular Polysaccharide Enables *Salmonella enterica* Serovar Typhi to Evade Microbe-Guided Neutrophil Chemotaxis. PLoS Pathog **10**:e1004306–9.

55. **Wilson RP**, **Raffatellu M**, **Chessa D**, **Winter SE**, **Tükel C**, **Bäumler AJ**. 2008. The Vi-capsule prevents Toll-like receptor 4 recognition of *Salmonella*. Cellular Microbiology **10**:876–890.

56. **Zhang X-L**, **Jeza VT**, **Pan Q**. 2008. *Salmonella* typhi: from a human pathogen to a vaccine vector. Cell Mol Immunol **5**:91–97.

57. **Anderson RM**, **May RM**. 1982. Coevolution of hosts and parasites. Parasitology **85 (Pt 2)**:411–426.

58. **Baumler A**, **Fang FC**. 2013. Host Specificity of Bacterial Pathogens. Cold Spring Harbor Perspectives in Medicine **3**:a010041–a010041.

59. **Woolhouse ME**, **Taylor LH**, **Haydon DT**. 2001. Population biology of multihost pathogens. Science **292**:1109–1112.

60. **Blancou J**, **Aubert MF**. 1997. [Transmission of rabies virus: importance of the species barrier]. Bulletin de l'Academie nationale de medecine **181**:301–11– discussion 311–2.

61. **Garnick E**. 1992. Parasite virulence and parasite-host coevolution: a reappraisal. J Parasitol **78**:381–386.

62. **Taylor LH**, **Latham SM**, **Woolhouse ME**. 2001. Risk factors for human disease emergence. Philosophical Transactions of the Royal Society B: Biological Sciences **356**:983–989.

63. **Cleaveland S**, **Laurenson MK**, **Taylor LH**. 2001. Diseases of humans and their domestic mammals: pathogen characteristics, host range and the risk of emergence. Philosophical Transactions of the Royal Society B: Biological Sciences **356**:991–999.

64. **Shivaprasad HL**. 2000. Fowl typhoid and pullorum disease. Rev Sci Tech **19**:405–424.

65. **Schiøler H**, **Christiansen ED**, **Høybye G**. 1983. Biliary calculi in chronic *Salmonella* carriers and healthy controls: a controlled study. Scandinavian Journal of Infectious Diseases **15**:17–19.

66. **Marshall JM**, **Flechtner AD**, **La Perle KM**, **Gunn JS**. 2014. Visualization of Extracellular Matrix Components within Sectioned *Salmonella* Biofilms on the Surface of Human Gallstones. PloS one **9**:e89243–9.

67. **Kozlica J**, **Claudet AL**, **Solomon D**, **Dunn JR**, **Carpenter LR**. 2010. Waterborne Outbreak of *Salmonella* 4,[5],12:i:. Foodborne Pathogens and Disease **7**:1431–1433.

68. **Berg R**. 2008. The Alamosa *Salmonella* outbreak: a gumshoe investigation. J Environ Health **71**:54–58.

69. **Schuster CJ**, **Ellis AG**, **Robertson WJ**, **Charron DF**, **Aramini JJ**, **Marshall BJ**, **Medeiros DT**. 2005. Infectious disease outbreaks related to drinking water in Canada, 1974-2001. Can J Public Health **96**:254–258.

70. **Greene SK**, **Daly ER**, **Talbot EA**, **Demma LJ**, **Holzbauer S**, **Patel NJ**, **Hill TA**, **Walderhaug MO**, **Hoekstra RM**, **Lynch MF**, **Painter JA**. 2008. Recurrent multistate outbreak of *Salmonella* Newport associated with tomatoes from contaminated fields, 2005. Epidemiol Infect **136**:157–165.

71. **Centers for Disease Control and Prevention**. 2002. Multistate outbreaks of *Salmonella* serotype Poona infections associated with eating cantaloupe from Mexico--United States and Canada, 2000-2002. MMWR Morb Mortal Wkly Rep **51**:1044-1047.

72. **Centers for Disease Control and Prevention (CDC)**. 2008. Outbreak of *Salmonella* serotype Saintpaul infections associated with multiple raw produce items--United States, 2008. MMWR Morb Mortal Wkly Rep **57**:929–934.

73. **Baloda SB**, **Christensen L**, **Trajcevska S**. 2001. Persistence of a *Salmonella enterica* Serovar Typhimurium DT12 Clone in a Piggery and in Agricultural Soil Amended with *Salmonella*-Contaminated Slurry. Applied and Environmental Microbiology **67**:2859–2862.

74. **Collinson SK**, **Emödy L**, **Müller KH**, **Trust TJ**, **Kay WW**. 1991. Purification and characterization of thin, aggregative fimbriae from *Salmonella enteritidis*. Journal of Bacteriology **173**:4773–4781.

75. **Romling U**, **Sierralta WD**, **Eriksson K**, **Normark S**. 1998. Multicellular and aggregative behaviour of *Salmonella typhimurium* strains is controlled by mutations in the agfD promoter. Molecular Microbiology **28**:249–264.

76. **Solano C**, **García B**, **Valle J**, **Berasain C**, **Ghigo J-M**, **Gamazo C**, **Lasa I**. 2002. Genetic analysis of *Salmonella enteritidis* biofilm formation: critical role of cellulose. Molecular Microbiology **43**:793–808.

77. **Collinson SK**, **Doig PC**, **Doran JL**, **Clouthier S**, **Trust TJ**, **Kay WW**. 1993. Thin, aggregative fimbriae mediate binding of *Salmonella enteritidis* to fibronectin. Journal of Bacteriology **175**:12–18.

78. **Gibson D**, **White A**, **Snyder S**, **Martin S**, **Heiss C**, **Azadi P**, **Surette M**, **Kay W**. 2006. *Salmonella* produces an O-antigen capsule regulated by AgfD and important for environmental persistence. Journal of Bacteriology **188**:7722.

79. **Zogaj X**, **Nimtz M**, **Rohde M**, **Bokranz W**, **Romling U**. 2001. The multicellular morphotypes of *Salmonella typhimurium* and *Escherichia coli* produce cellulose as the second component of the extracellular matrix. Molecular Microbiology **39**:1452–1463.

80. **White AP**, **Gibson DL**, **Kim W**, **Kay WW**, **Surette MG**. 2006. Thin aggregative fimbriae and cellulose enhance long-term survival and persistence of *Salmonella*. Journal of Bacteriology **188**:3219–3227.

81. **Anriany YA**, **Weiner RM**, **Johnson JA**, **De Rezende CE**, **Joseph SW**. 2001. *Salmonella enterica* Serovar Typhimurium DT104 Displays a Rugose Phenotype. Applied and Environmental Microbiology **67**:4048–4056.

82. **Scher K**, **Romling U**, **Yaron S**. 2005. Effect of Heat, Acidification, and Chlorination on *Salmonella enterica* Serovar Typhimurium Cells in a Biofilm Formed at the Air-Liquid Interface. Applied and Environmental Microbiology **71**:1163–1168.

83. **Tabak M**, **Scher K**, **Chikindas ML**, **Yaron S**. 2009. The synergistic activity of

triclosan and ciprofloxacin on biofilms of *Salmonella* Typhimurium. FEMS Microbiology Letters **301**:69–76.

84. **Apel D**, **White A**, **Grassl G**, **Finlay B**, **Surette M**. 2009. Long Term Survival of *Salmonella enterica* Serovar Typhimurium reveals an infectious state that is underrepresented on laboratory media containing bile salts. Applied and Environmental Microbiology. **75**:4923-4925.

85. **White A**, **Gibson D**, **Grassl G**, **Kay W**, **Finlay B**, **Vallance B**, **Surette M**. 2008. Aggregation via the red, dry, and rough morphotype is not a virulence adaptation in *Salmonella enterica* serovar Typhimurium. Infection and Immunity **76**:1048-1058.

86. **Barnhart MM**, **Chapman MR**. 2006. Curli biogenesis and function. Annu Rev Microbiol **60**:131–147.

87. **Olsén A**, **Jonsson A**, **Normark S**. 1989. Fibronectin binding mediated by a novel class of surface organelles on *Escherichia coli*. Nature **338**:652–655.

88. **Collinson SK**, **Clouthier SC**, **Doran JL**, **Banser PA**, **Kay WW**. 1996. *Salmonella enteritidis* agfBAC operon encoding thin, aggregative fimbriae. Journal of Bacteriology **178**:662–667.

89. **Romling U**, **Bian Z**, **Hammar M**, **Sierralta WD**, **Normark S**. 1998. Curli fibers are highly conserved between *Salmonella typhimurium* and *Escherichia coli* with respect to operon structure and regulation. Journal of Bacteriology **180**:722–731.

90. **White AP**, **Collinson SK**, **Banser PA**, **Gibson DL**, **Paetzel M**, **Strynadka NCJ**, **Kay WW**. 2001. Structure and characterization of AgfB from *Salmonella enteritidis* thin aggregative fimbriae. Journal of Molecular Biology **311**:735–749.

91. **Hammar M**, **Arnqvist A**, **Bian Z**, **Olsén A**, **Normark S**. 1995. Expression of two csg operons is required for production of fibronectin- and congo red-binding curli polymers in *Escherichia coli* K-12. Molecular Microbiology **18**:661–670.

92. **Olsén A**, **Arnqvist A**, **Hammar M**, **Sukupolvi S**, **Normark S**. 1993. The RpoS sigma factor relieves H-NS-mediated transcriptional repression of csgA, the subunit gene of fibronectin-binding curli in *Escherichia coli*. Molecular Microbiology **7**:523–536.

93. **Gerstel U**, **Park C**, **Romling U**. 2003. Complex regulation of csgD promoter activity by global regulatory proteins. Molecular Microbiology **49**:639–654.

94. **Arnqvist A**, **Olsén A**, **Normark S**. 1994. σS-dependent growth-phase induction of the csgBA promoter in *Escherichia coli* can be achieved in vivo by σ70 in the absence of the nucleoid-associated protein H-NS. Molecular Microbiology

13:1021–1032.

95. **Robbe-Saule V**, **Jaumouillé V**, **Prévost M**, **Guadagnini S**, **Talhouarne C**, **Mathout H**, **Kolb A**, **Norel F**. 2006. Crl activates transcription initiation of RpoS-regulated genes involved in the multicellular behavior of *Salmonella enterica* serovar Typhimurium. Journal of Bacteriology **188**:3983.

96. **Prigent-Combaret C**, **Brombacher E**, **Vidal O**, **Ambert A**, **Lejeune P**, **Landini P**, **Dorel C**. 2001. Complex Regulatory Network Controls Initial Adhesion and Biofilm Formation in *Escherichia coli* via Regulation of the csgD Gene. Journal of Bacteriology **183**:7213–7223.

97. **Steenackers H**, **Hermans K**, **Vanderleyden J**, **De Keersmaecker SCJ**. 2011. *Salmonella* biofilms: An overview on occurrence, structure, regulation and eradication. FRIN 1–30.

98. **Gerstel U**, **Kolb A**, **Romling U**. 2006. Regulatory components at the csgD promoter - additional roles for OmpR and integration host factor and role of the 5' untranslated region. FEMS Microbiology Letters **261**:109–117.

99. **Ogasawara H**, **Yamamoto K**, **Ishihama A**. 2010. Regulatory role of MlrA in transcription activation of csgD, the master regulator of biofilm formation in *Escherichia coli*. FEMS Microbiology Letters **312**:160–168.

100. **Brown PK**, **Dozois CM**, **Nickerson CA**, **Zuppardo A**, **Terlonge J**, **Curtiss R**. 2001. MlrA, a novel regulator of curli (AgF) and extracellular matrix synthesis by *Escherichia coli* and *Salmonella enterica* serovar Typhimurium. Molecular Microbiology **41**:349–363.

101. **Holmqvist E**, **rd JRA**, **Sterk M**, **Grantcharova N**, **mling URO**, **Wagner EGH**. 2010. Two antisense RNAs target the transcriptional regulator CsgD to inhibit curli synthesis. The EMBO Journal **29**:1840–1850.

102. **Jorgensen MG**, **Nielsen JS**, **Boysen A**, **Franch T**, **Moller-Jensen J**, **Valentin-Hansen P**. 2012. Small regulatory RNAs control the multi-cellular adhesive lifestyle of *Escherichia coli*. Molecular Microbiology **84**:36–50.

103. **Mika F**, **Busse S**, **Possling A**, **Berkholz J**, **Tschowri N**, **Sommerfeldt N**, **Pruteanu M**, **Hengge R**. 2012. Targeting of csgD by the small regulatory RNA RprA links stationary phase, biofilm formation and cell envelope stress in *Escherichia coli*. Molecular Microbiology **84**:51–65.

104. **Thomason MK**, **Fontaine F**, **De Lay N**, **Storz G**. 2012. A small RNA that regulates motility and biofilm formation in response to changes in nutrient availability in *Escherichia coli.* Molecular Microbiology **84**:17–35.

105. **Serra DO**, **Mika F**, **Richter AM**, **Hengge R**. 2016. The green tea polyphenol EGCG inhibits *E. coli* biofilm formation by impairing amyloid curli fibre

assembly and downregulating the biofilm regulator CsgD via the sigma(E) - dependent sRNA RybB. Molecular Microbiology **101**:136–151.

106. **Boehm A**, **Vogel J**. 2012. The csgD mRNA as a hub for signal integration via multiple small RNAs. Molecular Microbiology **84**:1–5.

107. **Mika F**, **Hengge R**. 2014. Small RNAs in the control of RpoS, CsgD, and biofilm architecture of *Escherichia coli*. RNA Biol **11**:494–507.

108. **Romling U**, **Rohde M**, **Olsén A**, **Normark S**, **Reinköster J**. 2000. AgfD, the checkpoint of multicellular and aggregative behaviour in *Salmonella typhimurium* regulates at least two independent pathways. Molecular Microbiology **36**:10–23.

109. **Simm R**, **Morr M**, **Kader A**, **Nimtz M**, **Römling U**. 2004. GGDEF and EAL domains inversely regulate cyclic di-GMP levels and transition from sessility to motility. Molecular Microbiology **53**:1123–1134.

110. **Römling U**, **Galperin MY**. 2015. Bacterial cellulose biosynthesis: diversity of operons, subunits, products, and functions. Trends Microbiol **23**:545–557.

111. **Tal R**, **Wong HC**, **Calhoon R**, **Gelfand D**, **Fear AL**, **Volman G**, **Mayer R**, **Ross P**, **Amikam D**, **Weinhouse H**, **Cohen A**, **Sapir S**, **Ohana P**, **Benziman M**. 1998. Three cdg operons control cellular turnover of cyclic di-GMP in *Acetobacter xylinum*: genetic organization and occurrence of conserved domains in isoenzymes. Journal of Bacteriology **180**:4416–4425.

112. **Ross P**, **Weinhouse H**, **Aloni Y**, **Michaeli D**, **Weinberger-Ohana P**, **Mayer R**, **Braun S**, **de Vroom E**, **van der Marel GA**, **van Boom JH**, **Benziman M**. 1987. Regulation of cellulose synthesis in *Acetobacter xylinum* by cyclic diguanylic acid. Nature **325**:279–281.

113. **Romling U**, **Galperin MY**, **Gomelsky M**. 2013. Cyclic di-GMP: the First 25 Years of a Universal Bacterial Second Messenger. Microbiology and Molecular Biology Reviews **77**:1–52.

114. **Simm R**, **Lusch A**, **Kader A**, **Andersson M**, **Römling U**. 2007. Role of EAL-containing proteins in multicellular behavior of *Salmonella enterica* serovar Typhimurium. Journal of Bacteriology **189**:3613–3623.

115. **García B**, **Latasa C**, **Solano C**, **Portillo FG-D**, **Gamazo C**, **Lasa I**. 2004. Role of the GGDEF protein family in *Salmonella* cellulose biosynthesis and biofilm formation. Molecular Microbiology **54**:264–277.

116. **Solano C**, **García B**, **Latasa C**, **Toledo-Arana A**, **Zorraquino V**, **Valle J**, **Casals J**, **Pedroso E**, **Lasa I**, **Greenberg EP**. 2009. Genetic Reductionist Approach for Dissecting Individual Roles of GGDEF Proteins within the c-di-GMP Signaling Network in *Salmonella*. Proceedings of the National Academy of

Sciences of the United States of America **106**:7997–8002.

117. **Kader A**, **Simm R**, **Gerstel U**, **Morr M**, **Römling U**. 2006. Hierarchical involvement of various GGDEF domain proteins in rdar morphotype development of *Salmonella enterica* serovar Typhimurium. Molecular Microbiology **60**:602–616.

118. **Latasa C**, **Roux A**, **Toledo-Arana A**, **Ghigo J-M**, **Gamazo C**, **Penadés JR**, **Lasa I**. 2005. BapA, a large secreted protein required for biofilm formation and host colonization of *Salmonella enterica* serovar Enteritidis. Molecular Microbiology **58**:1322–1339.

119. **Cucarella C**, **Solano C**, **Valle J**, **Amorena B**, **Lasa I**, **Penades JR**. 2001. Bap, a *Staphylococcus aureus* surface protein involved in biofilm formation. Journal of Bacteriology **183**:2888–2896.

120. **Jonas K**, **Tomenius H**, **Kader A**, **Normark S**, **Römling U**, **Belova LM**, **Melefors Ö**. 2007. Roles of curli, cellulose and BapA in *Salmonella* biofilm morphology studied by atomic force microscopy. BMC Microbiology **7**:70–9.

121. **White AP**, **Gibson DL**, **Collinson SK**, **Banser PA**, **Kay WW**. 2003. Extracellular Polysaccharides Associated with Thin Aggregative Fimbriae of *Salmonella enterica* Serovar Enteritidis. Journal of Bacteriology **185**:5398–5407.

122. **Snyder DS**, **Gibson D**, **Heiss C**, **Kay W**, **Azadi P**. 2006. Structure of a capsular polysaccharide isolated from *Salmonella enteritidis*. Carbohydr Res **341**:2388–2397.

123. **Barak JD**, **Jahn CE**, **Gibson DL**, **Charkowski AO**. 2007. The role of cellulose and O-antigen capsule in the colonization of plants by *Salmonella enterica*. Mol Plant Microbe Interact **20**:1083–1091.

124. **Römling U**, **Bokranz W**, **Rabsch W**, **Zogaj X**, **Nimtz M**, **Tschäpe H**. 2003. Occurrence and regulation of the multicellular morphotype in *Salmonella* serovars important in human disease. International Journal of Medical Microbiology **293**:273–285.

125. **Solomon EB**, **Niemira BA**, **Sapers GM**, **Annous BA**. 2005. Biofilm formation, cellulose production, and curli biosynthesis by *Salmonella* originating from produce, animal, and clinical sources. J Food Prot **68**:906–912.

126. **White A**, **Surette M**. 2006. Comparative genetics of the rdar morphotype in *Salmonella*. Journal of Bacteriology **188**:8395.

127. **Singletary LA**, **Karlinsey JE**, **Libby SJ**, **Mooney JP**, **Lokken KL**, **Tsolis RM**, **Byndloss MX**, **Hirao LA**, **Gaulke CA**, **Crawford RW**, **Dandekar S**, **Kingsley RA**, **Msefula CL**, **Heyderman RS**, **Fang FC**. 2016. Loss of Multicellular Behavior in Epidemic African Nontyphoidal *Salmonella enterica* Serovar

Typhimurium ST313 Strain D23580. mBio **7**:e02265–15–11.

128. **Baumler AJ**, **Gilde AJ**, **Tsolis RM**, **van der Velden AW**, **Ahmer BM**, **Heffron F**. 1997. Contribution of horizontal gene transfer and deletion events to development of distinctive patterns of fimbrial operons during evolution of *Salmonella* serotypes. Journal of Bacteriology **179**:317–322.

129. **Doran JL**, **Collinson SK**, **Burian J**, **Sarlós G**, **Todd EC**, **Munro CK**, **Kay CM**, **Banser PA**, **Peterkin PI**, **Kay WW**. 1993. DNA-based diagnostic tests for *Salmonella* species targeting agfA, the structural gene for thin, aggregative fimbriae. Journal of Clinical Microbiology **31**:2263–2273.

130. **Zogaj X**, **Bokranz W**, **Nimtz M**, **Romling U**. 2003. Production of Cellulose and Curli Fimbriae by Members of the Family *Enterobacteriaceae* Isolated from the Human Gastrointestinal Tract. Infection and Immunity **71**:4151–4158.

131. **Bokranz W**, **Wang X**, **Tschäpe H**, **Römling U**. 2005. Expression of cellulose and curli fimbriae by *Escherichia coli* isolated from the gastrointestinal tract. Journal of Medical Microbiology **54**:1171–1182.

132. **Fux CA**, **Costerton JW**, **Stewart PS**, **Stoodley P**. 2005. Survival strategies of infectious biofilms. Trends Microbiol **13**:34–40.

133. **Davidson CJ**, **White AP**, **Surette MG**. 2008. Evolutionary loss of the rdar morphotype in *Salmonella* as a result of high mutation rates during laboratory passage. ISME J **2**:293–307.

134. **Behr MA**, **Wilson MA**, **Gill WP**, **Salamon H**, **Schoolnik GK**, **Rane S**, **Small PM**. 1999. Comparative genomics of BCG vaccines by whole-genome DNA microarray. Science **284**:1520–1523.

135. **Uhlich GA**, **Keen JE**, **Elder RO**. 2001. Mutations in the csgD promoter associated with variations in curli expression in certain strains of *Escherichia coli* O157:H7. Applied and Environmental Microbiology **67**:2367–2370.

136. **Parkhill J**, **Dougan G**, **James KD**, **Thomson NR**, **Pickard D**, **Wain J**, **Churcher C**, **Mungall KL**, **Bentley SD**, **Holden MT**, **Sebaihia M**, **Baker S**, **Basham D**, **Brooks K**, **Chillingworth T**, **Connerton P**, **Cronin A**, **Davis P**, **Davies RM**, **Dowd L**, **White N**, **Farrar J**, **Feltwell T**, **Hamlin N**, **Haque A**, **Hien TT**, **Holroyd S**, **Jagels K**, **Krogh A**, **Larsen TS**, **Leather S**, **Moule S**, **O'Gaora P**, **Parry C**, **Quail M**, **Rutherford K**, **Simmonds M**, **Skelton J**, **Stevens K**, **Whitehead S**, **Barrell BG**. 2001. Complete genome sequence of a multiple drug resistant *Salmonella enterica* serovar Typhi CT18. Nature **413**:848–852.

137. **White A**, **Weljie A**, **Apel D**, **Zhang P**. 2010. A Global Metabolic Shift Is Linked to *Salmonella* Multicellular Development. PloS one. **5**:e11814.

138. **Grantcharova N**, **Peters V**, **Monteiro C**, **Zakikhany K**, **Romling U**. 2010. Bistable expression of CsgD in biofilm development of *Salmonella enterica* serovar Typhimurium. Journal of Bacteriology **192**:456.

139. **Dubnau D**, **Losick R**. 2006. Bistability in bacteria. Molecular Microbiology **61**:564–572.

140. **Gualdi L**, **Tagliabue L**, **Landini P**. 2007. Biofilm formation-gene expression relay system in *Escherichia coli*: modulation of sigmaS-dependent gene expression by the CsgD regulatory protein via sigmaS protein stabilization. Journal of Bacteriology **189**:8034–8043.

141. **Bougdour A**, **Wickner S**, **Gottesman S**. 2006. Modulating RssB activity: IraP, a novel regulator of σS stability in *Escherichia coli*. Genes Dev.

142. **Brombacher E**, **Baratto A**, **Dorel C**, **Landini P**. 2006. Gene expression regulation by the Curli activator CsgD protein: modulation of cellulose biosynthesis and control of negative determinants for microbial adhesion. Journal of Bacteriology **188**:2027–2037.

143. **Piégu B**, **Bire S**, **Arensburger P**, **Bigot Y**. 2015. A survey of transposable element classification systems--a call for a fundamental update to meet the challenge of their diversity and complexity. Mol Phylogenet Evol **86**:90–109.

144. **Finnegan DJ**. 1989. Eukaryotic transposable elements and genome evolution. Trends in Genetics **5 IS -**:103–107.

145. **Barth PT**, **Datta N**, **Hedges RW**. 1976. Transposition of a deoxyribonucleic acid sequence encoding trimethoprim and streptomycin resistances from R483 to other replicons. Journal of Bacteriology.

146. **Waddell CS**, **Craig NL**. 1988. Tn7 transposition: two transposition pathways directed by five Tn7-encoded genes. Genes Dev **2**:137–149.

147. **Hedges RW**, **Datta N**, **Fleming MP**. 1972. R factors conferring resistance to trimethoprim but not sulphonamides. J Gen Microbiol **73**:573–575.

148. **Parks AR**, **Peters JE**. 2007. Transposon Tn7 Is Widespread in Diverse Bacteria and Forms Genomic Islands. Journal of Bacteriology **189**:2170–2173.

149. **Peters JE**. 2014. Tn7. Microbiology Spectrum **2**:1–20.

150. **McKenzie GJ**, **Craig NL**. 2006. Fast, easy and efficient: site-specific insertion of transgenes into Enterobacterial chromosomes using Tn*7* without need for selection of the insertion event. BMC Microbiology **6**:39.

151. **Bainton RJ**, **Kubo KM**, **Feng JN**, **Craig NL**. 1993. Tn7 transposition: target DNA recognition is mediated by multiple Tn7-encoded proteins in a purified in

vitro system. Cell **72**:931–943.

152. **Sharpe PL**, **Craig NL**. 1998. Host proteins can stimulate Tn7 transposition: a novel role for the ribosomal protein L29 and the acyl carrier protein. The EMBO Journal **17**:5822–5831.

153. **Kuduvalli PN**, **Rao JE**, **Craig NL**. 2001. Target DNA structure plays a critical role in Tn7 transposition. The EMBO Journal **20**:924–932.

154. **Lichtenstein C**, **Brenner S**. 1982. Unique insertion site of Tn7 in the *E. coli* chromosome. Nature **297**:601–603.

155. **DeBoy RT**, **Craig NL**. 1996. Tn7 transposition as a probe of cis interactions between widely separated (190 kilobases apart) DNA sites in the *Escherichia coli* chromosome. J Bacteriol **178**:6184-6191.

156. **Bainton R**, **Gamas P**, **Craig NL**. 1991. Tn7 transposition in vitro proceeds through an excised transposon intermediate generated by staggered breaks in DNA. Cell **65**:805–816.

157. **Ronning DR**, **Li Y**, **Perez ZN**, **Ross PD**, **Hickman, AB**, **Craig NL**, **Dyda F**. 2004. The carboxy-terminal portion of TnsC activates the Tn7 transposase through a specific interaction with TnsA. The EMBO Journal **23**:2972–2981.

158. **May EW**, **Craig NL**. 1996. Switching from cut-and-paste to replicative Tn7 transposition. Science **272**:401–404.

159. **Sarnovsky RJ**, **May EW**, **Craig NL**. 1996. The Tn7 transposase is a heteromeric complex in which DNA breakage and joining activities are distributed between different gene products. The EMBO Journal **15**:6348–6361.

160. **Arciszewska LK**, **Craig NL**. 1991. Interaction of the Tn7-encoded transposition protein TnsB with the ends of the transposon. Nucleic Acids Research **19**:5021–5029.

161. **Arciszewska LK**, **McKown RL**, **Craig NL**. 1991. Purification of TnsB, a transposition protein that binds to the ends of Tn7. J Biol Chem **266**:21736–21744.

162. **Parks AR**, **Peters JE**. 2009. Tn7 elements: Engendering diversity from chromosomes to episomes. Plasmid **61**:1–14.

163. **Choi, K. Y., Li, Y., Sarnovsky, R., & Craig, N. L.** (2013). Direct interaction between the TnsA and TnsB subunits controls the heteromeric Tn7 transposase. *PNAS* **110**:E2038–45. http://doi.org/10.1073/pnas.1305716110.

164. **Peters JE**, **Craig NL**. 2001. Tn7 recognizes transposition target structures associated with DNA replication using the DNA-binding protein TnsE. Genes

Dev **15**:737–747.

165. **Parks AR**, **Li Z**, **Shi Q**, **Owens RM**, **Jin MM**, **Peters JE**. 2009. Transposition into Replicating DNA Occurs through Interaction with the Processivity Factor. Cell **138**:685–695.

166. **Wolkow CA**, **DeBoy RT**, **Craig NL**. 1996. Conjugating plasmids are preferred targets for Tn7. Genes Dev **10**:2145–2157.

167. **Lawley T**, **Frost LS**, **Wilkins BM**. 2004. Bacterial Conjugation in Gram-Negative Bacteria†, pp. 203–226. *In* Plasmid biology. American Society of Microbiology.

168. **Crepin S**, **Harel J**, **Dozois CM**. 2012. Chromosomal Complementation Using Tn7 Transposon Vectors in Enterobacteriaceae. Applied and Environmental Microbiology **78**:6001–6008.

169. **Knodler LA**, **Bestor A**, **Ma C**, **Hansen-Wester I**, **Hensel M**, **Vallance BA**, **Steele-Mortimer O**. 2005. Cloning Vectors and Fluorescent Proteins Can Significantly Inhibit *Salmonella* enterica Virulence in Both Epithelial Cells and Macrophages: Implications for Bacterial Pathogenesis Studies. Infection and Immunity **73**:7027–7031.

170. **Choi K-H**, **Gaynor JB**, **White KG**, **Lopez C**, **Bosio CM**, **Karkhoff-Schweizer RR**, **Schweizer HP**. 2005. A Tn7-based broad-range bacterial cloning and expression system. Nat Meth **2**:443–448.

171. **Bumann D**, **Hueck C**, **Aebischer T**, **Meyer TF**. 2000. Recombinant live Salmonellaspp. for human vaccination against heterologous pathogens. FEMS Immunol Med Microbiol **27**:357–364.

172. **Bao Y**, **Lies DP**, **Fu H**, **Roberts GP**. 1991. An improved Tn7-based system for the single-copy insertion of cloned genes into chromosomes of gram-negative bacteria. Gene **109**:167–168.

173. **Koch B**, **Jensen LE**, **Nybroe O**. 2001. A panel of Tn7-based vectors for insertion of the gfp marker gene or for delivery of cloned DNA into Gram-negative bacteria at a neutral chromosomal site. Journal of Microbiological Methods **45**:187–195.

174. **Staley TE**, **Lawrence EG**, **Drahos DJ**. 1997. Variable specificity of Tn7:: lacZY insertion into the chromosome of root-colonizing Pseudomonas putida strains. Molecular Ecology **6**:85–87.

175. **McKenzie GJ**, **Craig NL**. 2006. Fast, easy and efficient: site-specific insertion of transgenes into Enterobacterial chromosomes using Tn*7* without need for selection of the insertion event. BMC Microbiology **6**:39–7.

176. **Højberg O**, **Schnider U**, **Winteler HV**, **Sørensen J**, **Haas D**. 1999. Oxygen-sensing reporter strain of *Pseudomonas fluorescens* for monitoring the distribution of low-oxygen habitats in soil. Applied and Environmental Microbiology **65**:4085–4093.

177. **Choi KH**, **Mima T**, **Casart Y**, **Rholl D**, **Kumar A**, **Beacham IR**, **Schweizer HP**. 2008. Genetic Tools for Select-Agent-Compliant Manipulation of *Burkholderia pseudomallei*. Applied and Environmental Microbiology **74**:1064–1075.

178. **Shivak DJ**, **MacKenzie KD**, **Watson N**, **Pasternak A**, **Jones BD**, **Wang Y**, **DeVinney R**, **Wilson H**, **Surette MG**, **White AP**. 2016. A modular, Tn7-based system for making bioluminescent or fluorescent *Salmonella* and *E. coli* strains. Applied and Environmental Microbiology AEM.01346–16–15.

179. **Stewart, P. S., & Franklin, M. J.** 2008. Physiological heterogeneity in biofilms. Nature Reviews. Microbiology. **6**:199–210.

180. **Finn, J. A., Parks, A. R., & Peters, J. E.** 2007. Transposon Tn7 directs transposition into the genome of filamentous bacteriophage M13 using the element-encoded TnsE protein. Journal of Bacteriology. **189**:9122–9125.

2.0 RATIONALE, HYPOTHESES, AND OBJECTIVES

2.1 Rationale and Hypotheses

Salmonella biofilm formation was traditionally studied at the population level as the rdar morphotype. The rdar morphotype was activated by environmental conditions, such as nutrient limitation, low temperatures, and low osmolarity, and was shown to enhance the survival and persistence of *Salmonella* cells. Cells within a biofilm retained the ability to cause an infection, even after exposure to environmental conditions (i.e. desiccation and starvation) for a long period of time – on the order of years. However, it was still viewed primarily as a survival adaptation, and not a virulence trait. A comparison of biofilm-positive and biofilm-negative colonies revealed a shift in the expression of multiple genes associated with metabolism during biofilm formation. Many of these changes did not appear to have a direct relationship to CsgD, which suggested that biofilm formation could represent a developmental program.

The paradigm of *Salmonella* biofilm formation as a population-level phenotype was shifted by the discovery that CsgD expression varies between cells. The ON-OFF nature of CsgD bistability results in two different cell types – multicellular aggregates and planktonic cells. These subpopulations are genetically clonal, but are distinct at the phenotypic level. Thus, in my Ph.D. research we set out to perform a detailed comparison between multicellular aggregates and planktonic cells to better understand the potential role of CsgD bistability in *Salmonella* biology. We hypothesized that multicellular aggregates would have a similar gene expression profile to *Salmonella* rdar morphotype cells. In contrast, we predicted that planktonic cells would have a novel gene expression profile (i.e., their gene expression would not match the expression that has been characterized for *Salmonella* rdar morphotype cells).

For the last section of my thesis, we investigated the biofilm-forming capacity of invasive nontyphoidal *Salmonella* isolates. Most nontyphoidal *Salmonella* isolates are associated with gastroenteritis, are transiently associated with the human host, and are positive for the biofilm phenotype. In contrast, host-restricted typhoidal *Salmonella* isolates are associated with an invasive systemic infection, can persist asymptomatically within humans, and are biofilm-negative. The invasive nontyphoidal isolates represent a

phylogenetically distinct set of nontyphoidal *Salmonella* strains that are associated with a typhoid-like illness, primarily in sub-Saharan Africa. Invasive NTS strains can be isolated from asymptomatic persons, and have yet to be isolated from an environmental reservoir. Due to their apparent similarities to typhoidal strains, we hypothesized that invasive nontyphoidal *Salmonella* strains would be negative for the biofilm phenotype.

2.2 Objectives

The original objectives of my research project included the following:

i. Generate a *csgD*-complemented *Salmonella* strain for use as a control strain in experiments (Chapter 3).
ii. Develop a method for efficient chromosomal integration in *Salmonella* (Chapter 3).
iii. Use *csgD* and *csgB* promoter expression to identify time points corresponding to important stages of *Salmonella* biofilm development during growth in liquid culture (Chapter 4).
iv. Develop a procedure for separating multicellular aggregates from planktonic cells (Chapter 4).
v. Develop a method for reproducible RNA extraction from multicellular aggregates and planktonic cells (Chapter 4).
vi. Develop a method for the robust quantitation and enumeration of cells within multicellular aggregates (Chapter 4).

Based on the results from our experiments, additional objectives were defined:

vii. Determine if increased SPI-1 and SPI-2 type three secretion system expression in planktonic cells was reflected in corresponding increases in protein synthesis (Chapter 4).
viii. Compare the virulence and survival properties of multicellular aggregates and planktonic cells (Chapter 4).
ix. Evaluate iNTS strains for the ability to form a biofilm phenotype (Chapter 5).
x. Evaluate the biofilm regulatory network in iNTS strains compared to known biofilm-positive and biofilm-negative *Salmonella* strains (Chapter 5).

3.0 A MODULAR, TN7-BASED SYSTEM FOR MAKING BIOLUMINESCENT OR FLUORESCENT *SALMONELLA* AND *E. COLI* STRAINS

Dylan J. Shivak[1,2]†, Keith D. MacKenzie[1,2]†, Nikole Watson[1,2], Alex Pasternak[1], Brian D. Jones[3], Yejun Wang[4], Rebekah DeVinney[3], Heather Wilson[1], Michael G. Surette[5] and Aaron P. White[1,2]*

[1] Vaccine and Infectious Disease Organization-International Vaccine Centre, Saskatoon, Saskatchewan, Canada

[2] Department of Microbiology and Immunology, University of Saskatchewan, Saskatoon, Saskatchewan, Canada

[3] Department of Microbiology, Immunology & Infectious Diseases, University of Calgary, Calgary, Alberta, Canada

[4] Department of Medical Genetics & Bioinformatics, Shenzhen University Health Science Center, Guangdong, China

[5] Farncombe Family Digestive Health Research Institute, McMaster University, Hamilton, Ontario, Canada

† Both authors contributed equally to this work.

* Corresponding author; e-mail: aaron.white@usask.ca

Full citation: Shivak, DJ, MacKenzie KD, Watson NL, Pasternak JA, Jones BD, Wang Y, DeVinney R, Wilson HL, Surette MG, White AP. (2016). A modular, Tn7-based system for making bioluminescent or fluorescent *Salmonella* and *E. coli* strains. Appl Environ Microbiol **82**(16):4931-4943.

Author contributions:

Keith MacKenzie and **Dylan Shivak** are attributed equal authorship, as they contributed equally to the data generation and manuscript preparation; as such some of this material was also presented in Dylan's Master of Science thesis "Exploring the dynamics of *Salmonella* transmission in a murine model of infection." (Dylan Shivak, 2015: Chapter 4, University of Saskatchewan, Saskatoon, SK).

Nikole Watson generated the GFP and mCherry promoter-reporter plasmids and strains and assisted with the experiment that evaluated the transposition efficiency of the Tn7 helper plasmids, pHSG-*tnsABCD*, pTNS2 and pTNS3.

Alex Pasternak performed the experiments for qRT-PCR with the Tn7 helper plasmids and confocal microscopy of GFP and mCherry reporter strains.

Brian Jones generated the chloramphenicol-resistant pCS26 plasmid.

Yejun Wang performed the bioinformatics analysis for the experiment evaluating the impact of the pCS26 plasmid on the transcriptome of *S.* Typhimurium 14028 cells.

Rebekah DeVinney, Heather Wilson, and **Michael Surette** provided guidance with experimental design and edited the manuscript.

Aaron White assisted with experimental design, performed *in vitro* experiments evaluating the GFP and mCherry reporter strains, assisted with statistical analyses, and held the research grant that supported this work.

3.1 Interface

This chapter presents the modifications we introduced to the existing Tn7 transposition system designed by Choi *et al.* (12). The genetic changes we introduced to this Tn7 system facilitated integration of antibiotic resistance genes and bioluminescent or fluorescent reporter markers into the chromosome of *Salmonella*. This tool for chromosomal tagging allowed us to track isogenic *Salmonella* strains (i.e., wildtype and Δ*csgD* strains of *Salmonella*) or *Salmonella* subpopulations (i.e., multicellular aggregates and planktonic cells). The marked *Salmonella* strains were used to validate the results of our transcriptomic analysis (presented in Chapter 4) by facilitating comparative physiological assays, such as *in vitro* fitness experiments or *in vivo* competitive infection assays.

3.2 Abstract

Our goal was to develop a robust tagging method that can be used to track bacterial strains *in vivo*. To address this challenge, we adapted two existing systems: a modular plasmid-based reporter system (pCS26) that has been used for high-throughput gene expression studies in *Salmonella* and *E. coli,* and Tn7 transposition. We generated kanamycin- and chloramphenicol-resistant versions of pCS26 with bacterial luciferase, GFP and mCherry reporters under the control of σ^{70}-dependent promoters to provide three different levels of constitutive expression. We improved upon the existing Tn7 system by modifying the delivery vector to accept pCS26 constructs and moving the transposase genes from a non-replicating helper plasmid into a temperature-sensitive plasmid that can be conditionally maintained. This resulted in a 10 to 30-fold boost in transposase gene expression and transposition efficiencies of 10^{-8} to 10^{-10} in *Salmonella* serovar Typhimurium and *E. coli* APEC O1, whereas the existing Tn7 system yielded no successful transposition events. The new reporter strains displayed reproducible signaling in microwell plate assays, confocal microscopy, and *in vivo* animal infections. We have combined two flexible and complementary tools that can be used for a multitude of molecular biology applications within the Enterobacteriaceae. This system can accommodate new promoter-reporter combinations as they become available and can

help to bridge the gap between modern, high-throughput technologies and classical molecular genetics.

3.3 Introduction

The ability to generate recombinant strains of *Escherichia coli* and *Salmonella enterica* has facilitated insight into their ecology and pathogenic mechanisms. In recent years, strain collections have been assembled consisting of single gene knockouts of the entire genomes of *E. coli* K-12 (1) and *S.* enterica serovar Typhimurium 14028 (2). These collections can be used for finely-tuned analysis of gene function, host-pathogen interactions, as well as for strain fitness and competition experiments (3). There are also a wide array of reporter systems available to analyze gene expression in detail, for the ordering of hierarchical gene circuits (4), a deeper understanding of metabolism (5) or the development of biosensors (6). The use of these systems, coupled with next generation sequencing approaches that have facilitated functional genomics (7, 8), allows for unprecedented access to and understanding of the inner workings of the microbial cell.

Despite these recent advances in molecular tools, there is a need for improved *in vivo* analysis of bacterial pathogens (9). It is necessary to validate the expression of individual genes identified by next generation sequencing applications, such as Tn-seq (10), or to complement *in vivo* – specific genetic techniques, such as IVET (11). Chromosomal integration is desirable because it provides improved stability and more stoichiometric expression as compared to plasmid-based expression. Since original publication of the miniTn7 transposition method in 2005 (12), it has proven to be an incredibly useful system for integrating genes into bacterial chromosomes, particularly with *Pseudomonas* species (13). Tn7 transposition is ideal because it is directional (14), highly-specific for *att*Tn7 sites (15), and results in single copy insertion in most bacterial species (13). For *Salmonella* and *E. coli*, the single *att*Tn7 site lies near the conserved *glmS* gene (16). Modifications have been made in an attempt to improve the flexibility of the Tn7 system, such as through the generation of new helper plasmids (17), helper plasmid and delivery plasmid hybrids (18), arabinose inducible gene expression (19), and luciferase integration (20, 21). Unfortunately, the Tn7 system (12, 18) did not work in our

hands for *Salmonella*. We speculated that this was due to low expression of the *tnsABCD* genes that are necessary for transposition to occur.

In this manuscript, we describe a modified miniTn7 system that is amenable for use in the Enterobacteriaceae and provides the means necessary to combine modern, high-throughput technologies with classical molecular genetics. We built a replication-proficient Tn7 helper plasmid based on a temperature-sensitive replicon and demonstrate that *tnsABCD* expression is boosted 10-30 times as compared to the original helper plasmids (12). Furthermore, we incorporated the pCS26 vector system, which was designed to be modular and has a long history of use in gene expression experiments (4, 22-24). This gives unprecedented control over promoter and reporter combinations, which can be tested *in vitro* and are only two steps away from chromosomal integration. We generated reporter strains of *S.* Typhimurium and *E. coli* with luminescent and fluorescent reporters and kanamycin or chloramphenicol antibiotic selection and describe synthetic σ^{70}-dependent promoters that provide constitutive expression at three different levels of expression. Further, we showcase the effectiveness of gene complementation using this system, an important requirement for the robust validation of gene function discovery. This system is amenable for use in the *Enterobacteriaceae* and provides the means necessary to combine classical molecular genetics with modern, high-throughput technologies in both *in vitro* and *in vivo* settings.

3.4 Materials and Methods

3.4.1 Bacterial strains, media, and growth conditions.

Salmonella enterica serovar Typhimurium strain ATCC 14028 (American Type Culture Collection, Manassas, VA) and *E. coli* APEC O1 (25) were used as the parental strains in this study. The isogenic *S.* Typhimurium Δ*csgD* strain, formerly named Δ*agfD*, has been described previously (26). All cloning experiments were performed in *E. coli* DH10B (pCS26, pHSG415 and pACYC184) or *E. coli* CC118 (lambda *pir*) (27) (pUC18R6K-mini-Tn7T). All oligonucleotides used are listed in Table 3.1. Prior to plasmid purification, cloning and analysis, strains were inoculated from frozen stocks onto Luria agar (Luria broth + 1.5% agar) supplemented with the appropriate antibiotics

(Kanamycin; Kn 50 μg mL^{-1}; Chloramphenicol; Cm 10 μg mL^{-1}) and grown for 20 h at 37˚C. One isolated colony was used to inoculate 5 mL LB broth cultures.

For bioluminescence assays, overnight cultures of *S.* Typhimurium or *E. coli* strains carrying pCS26-derivatives (*luxCDABE*, GFP or mCherry reporters) were diluted 1 in 600 in 1% tryptone broth or LB supplemented with 50 μg ml^{-1} Kn or 10 μg ml^{-1} Cm to a final volume of 150 mL in 96-well clear-bottom black plates (9520 Costar; Corning Inc.). pCS26 bacterial luciferase (*lux*) operon fusion plasmids containing *csgDEFG*, *csgBAC*, and *adrA* promoters have been described previously (26). The RpoS-dependent reporter plasmid, sig38H4, contains a promoter that was designed based on alignment of numerous RpoS-controlled promoters, driving expression of the *lux* operon (26, 28). For *E. coli* or *S.* Typhimurium strains with reporters integrated into the genome, final dilution was performed into media without antibiotics. The cultures in each well were overlaid with 50 mL of mineral oil prior to starting the assays. Cultures were assayed for absorbance (590-600 nm, 0.1s), luminescence (1s, counts per second (CPS)) or fluorescence (GFP; excitation 485nm, emission 535nm; mCherry: excitation 555 ±38 nm, emission 632 ±45 nm) every 30 min during growth at 37ºC with agitation in a Victor X3 multilabel plate reader (Perkin-Elmer).

Strains used in murine infection experiments were inoculated from frozen stocks onto Luria agar supplemented with 50 μg mL^{-1} Kn or 10 μg mL^{-1} Cm. Isolated colonies were selected and grown for 16 hours in LB broth at 37° C, shaking at 200 rpm. These cultures were used to generate the bacterial challenge solutions used to infect mice.

3.4.2 Generating a Cm^R pCS26 plasmid

The Kn-resistant (Kn^R) reporter plasmid pCS26 containing the promoter-less *luxCDABE* operon from *Photorhabdus luminescens* has been previously described (22). The *lux* gene products from *P. luminescens* are more thermostable than those from *Vibrio* species, and thus are more suitable for gene expression experiments (29, 30). To generate a chloramphenicol-resistant (Cm^R) version of pCS26, inverse PCR was performed using primers pCS26-Eco and pCS26-Pst to remove the *aph* gene coding for Kn^R. The linear PCR product was digested with *Eco*RI and *Pst*I (New England Biolabs). The *cat* gene coding for Cm^R was part of a 1.1 kb blunt-ended insert from a PCR Script AMP cloning

Table 3.1. Oligonucleotides used in this study.

Primer name	Sequence (5'-3')[a]	Purpose
pCS26-Eco	GCCGAATTCCGGCAAGAAAGCCATCCAGTTTACTT	To generate Cm^R-pCS26
pCS26-Pst	GCCCTGCAGGCGGGACTCTGGGGTTCGAGAGCTCGCTTGGACTCC	To generate Cm^R-pCS26
Bam_out_Pst	GCCCTGCAGGTCGACTCTAGACGATCCCC	To generate Cm^R-pCS26
Pst_to_terminator	CGGGACTCTGCAGTTCGAGAGCTCGCTTGGACTCCTG	To generate Cm^R-pCS26
sig70-16_sense	TCGAGAATAATTCTTTACATTTATGCTTCCGGCTCGTATTCTACGTGCAATTG	sig70_16 promoter
sig70-16_antisense	GATCCAATTGCACGTAGAATACGAGCCGGAAGCATAAATGTAAAGAATTATTC	
sig70_16-10c2F	TCGAGAATAATTCTTTACATTTATGCTTCCGGCTCGTATAATACGTGCAATTG	sig70c10 promoter
sig70_16-10c2R	GATCCAATTGCACGTATTATACGAGCCGGAAGCATAAATGTAAAGAATTATTC	
sig70_16-35c2F	TCGAGAATAATTCTTGACATTTATGCTTCCGGCTCGTATTCTACGTGCAATTG	sig70c35 promoter
sig70_16-35c2R	GATCCAATTGCACGTAGAATACGAGCCGGAAGCATAAATGTCAAGAATTATTC	
pZE05rev	TCACCGACAAACAACAGATA	Amplify mCherry
TNS2forEco	GATCGAATTCATGCTGTGAAAAAGCATACTGGACT	Amplify *tnsABCD* from pTNS2
TNS2revXma	GATCCCCGGGATTTTGGCTCGTTGAAGATCCGATG	
Sac_Pac_Kpn1	CTTAATTAAGGTAC	PacI polylinker inserted in pUC18R6K-mini-Tn7T
Sac_Pac_Kpn2	CTTAATTAAGAGCT	
mCherry-check	TGCGTGGTACTAATTTTCCA	To verify chromosomal insertion of mCherry
GFP-check	TACATCATGGCAGACAAACA	To verify chromosomal insertion of GFP
Lux-check	TCAACACTTGTTTCTTTGAGG	To verify chromosomal insertion of *luxCDABE*

glmSdetectFor	AACCACCCGTTCAGGCTGGCTA	To verify chromosomal insertion at *glmS* site
glmSdetectRev	ACGTTGACCAGCCGCGTAAC	
Cm-check	CCCCGTGGAGGTAATAATTG	To verify chromosomal insertion of Cm^R marker
Kn-check	TACCCGTGATATTGCTGAAG	To verify chromosomal insertion of Kn^R marker
csgDcompFor2	GATTCTCGGGCATCATAAATAACAATTTGT	To generate *csgD-cat* construct
csgDcompREV	GTTCGAATTCCCTTACCGCCTGAGATTATC	
csgDFOR_Bam	GGCCGGATCCCATCATAAATAACAATTTGT	
glmScsgDREV_Pst	GGCCCTGCAGGGCCGTCGATAGACGGCCTTTTTTTGTGCG CCGTGACAGGCGCTGTTCTTCCTTACCGCCTGAGATTATC	
glmSCAMfor_Eco	GGCCGAATTCAGCGCAGGTAGGCGTAGCACCTCTTAGTCG CTCTTCAGCCACCATAGAGAGTGTAGGCTGGAGCTGCTTC	
CAMrev_Bam	GGCCGGATCCATATGAATATCCTCCTTA	
CAMfor_Eco	GGCCGAATTCGTGTAGGCTGGAGCTGCTTC	
csgDREV_Kpn	GGCCGGTACCCCTTACCGCCTGAGATTATC	
tnsA-for	ATACGAGCCTTAACCGCAAA	qRT-PCR for *tnsA*
tnsA-rev	CGCAACTCCTCCATATTCAC	
tnsB-for	TGCCACGATTGCTGATATTT	qRT-PCR for *tnsB*
tnsB-rev	GCCACCACATAAGACGGATT	
tnsC-for	CGCATTGCTCGTTATTCTGA	qRT-PCR for *tnsC*
tnsC-rev	TGACGCTGATCTTCGGTATC	
tnsD-for	AATGCTAGGAGTCGCTGCTT	qRT-PCR for *tnsD*
tnsD-rev	TACCAATCTCGTTGCCAAAA	
fabG1	CAGCGTGCGGGTATCCTGGC	qRT-PCR – control gene
fabG2	CCGCCGTTGACGTGCAGAGT	
groEL1	CGTTGCGCTGATCCGCGTTG	qRT-PCR – control gene
groEL2	GACGGCTCTTCGCCGCAGTT	

[a] Nucleotide sequences corresponding to restriction enzyme sites are underlined.

kit (Agilent Technologies, Santa Clara, CA) that was cloned into *Sma*I-digested pUC18 and excised by digestion with *Eco*RI and *Pst*I. The inverse PCR product and the Cm^R fragment were ligated using T4 DNA ligase (New England Biolabs). To remove a second *Bam*HI site that was introduced as part of the Cm^R fragment, inverse PCR was performed using primers Bam_out_Pst and Pst_to_terminator. The product of inverse PCR was digested with *Pst*I and ligated using T4 DNA ligase to generate the final pCS26-Cm vector.

3.4.3 Generation of sig70_16, sig70c10 and sig70c35 promoter constructs

sig70_16 is a derivative of pCS26 that contains a synthetic, σ^{70}-dependent promoter controlling *luxCDABE* expression(31). To generate a sig70_16 version of pCS26-Cm^R, oligonucleotides sig70-16_sense and sig70-16_antisense were phosphorylated using T4 polynucleotide kinase (NEB) at 37ºC for 1 h in buffer containing 50mM Tris pH 8, 10mM $MgCl_2$, 5mM DTT, 1 mM ATP, and 1.5 mg of BSA. The phosphorylated oligonucleotides were mixed, heated to 80ºC for 7 min and allowed to cool slowly to 30ºC, and the annealed product was ligated into *Xho*I/*Bam*HI-digested pCS26-Cm^R. To generate sig70c10 and sig70c35 promoter variants, oligonucleotides sig70_16-10c2F and sig70_16-10c2R, and sig70_16-35c2F and sig70_16-35c2R were phosphorylated and annealed as above, then ligated into *Xho*I/*Bam*HI-digested pCS26-Kn^R or -Cm^R vectors. Clones were screened for light production using an IVIS Lumina II *in vivo* imaging system (Perkin Elmer). DNA sequencing (Eurofins MWG Operon) was performed on all constructs using primers pZE05 and pZE06 (22).

3.4.4 Generation of GFP and mCherry pCS26 reporter plasmids.

pCS26-Kn^R and -Cm^R containing sig70_16, sig70c10 and sig70c35 promoters were digested with *Not*I and the 7300 bp *luxCDABE*-containing fragment was removed. GFP-containing fragments were PCR amplified from pCS21 (22) and mCherry-containing fragments were PCR amplified from P*adrA*-mCherry (32) using primers pZE05 (22) and pZE05rev. Phusion High-Fidelity DNA polymerase (Fisher Scientific) was used and reaction conditions were as outlined by the manufacturer. The resulting products were digested with *Not*I, ligated into the pCS26 vectors and transformed into *E.*

coli DH10B. Potential clones were selected on Luria agar supplemented with the appropriate antibiotics and screened for fluorophore production on the IVIS Lumina II imaging system, as above. Proper orientation of reporter constructs was confirmed by plasmid purification followed by restriction digestion.

3.4.5 Modification of the Tn7 delivery and helper plasmids.

The *tnsABCD* operon was PCR-amplified from pTNS2 (12) with Phusion DNA polymerase using primers TNS2forEco and TNS2revXma. The *tnsABCD* PCR product was digested with *Eco*RI and *Xma*I and ligated into *Eco*RI/*Xma*I-digested pHSG415 (33). By cloning into this region of pHSG415, *tnsABCD* expression would be partly controlled by the existing *cat* promoter. To facilitate cloning of pCS26-based vector fragments into the pUC18R6K mini-Tn7T plasmid (12), a poly-linker containing a *Pac*I restriction site was inserted. Oligonucleotides Sac_Pac_Kpn1 and Sac_Pac_Kpn2 were phosphorylated and annealed as above and ligated into *Sac*I/*Kpn*I-digested pUC18R6K-mini-Tn7T. The pUC18R6K-mini-Tn7T-*Pac*I vector was used for all subsequent experiments.

3.4.6 Integration into the *E. coli* and *S.* Typhimurium chromosome.

pCS26-Kn^R and -Cm^R constructs were digested with *Pac*I to remove the pSC101-based origin of replication and ligated into *Pac*I-digested pUC16R6K mini-Tn7T *Pac*I vector. Potential clones in *E. coli* CC118 (lambda *pir*) were selected by growth on LB agar supplemented with 100 μg mL^{-1} Ap and 10 μg mL^{-1} Cm or 50 μg mL^{-1} Kn. To confirm that constructs were ligated in the desired orientation, plasmid purifications were performed, followed by restriction digestion. The resulting mini-Tn7T-pCS26 plasmids, 18 in total, were purified using Plasmid Midi kits (Qiagen). DNA sequencing was performed with primers pZE05 and pZE06 (Lux), pZE05 and mCherry-check (mCherry), or pZE05 and GFP-check (GFP) (Table 3.1).

Chromosomal integration was achieved by electroporation of 100 ng of the appropriate pUC18R6K-mini-Tn7T-pCS26 vector into competent cells of *S.* Typhimurium 14028 Δ*csgD* or *E. coli* APEC O1 containing pHSG415-*tnsABCD*. Competent cells were prepared by growing for 3-4 h at 28ºC, with agitation in LB supplemented with 100 μg mL^{-1} ampicillin. Cells were allowed to recover in SOC media

for 2 h at 30°C prior to plating on Luria agar supplemented with 50 μg mL^{-1} Kn or 10 μg mL^{-1} Cm. Final plates were incubated overnight at 37°C, to ensure loss of the temperature-sensitive pHSG415-*tnsABCD* plasmid. To test the existing Tn7 method, 100 ng of pTNS2 (12) or pTNS3 (34) were co-electroporated with 100 ng of miniTn7T-pCS26 plasmid into *S.* Typhimurium or *E. coli* competent cells. These plasmids do not replicate in *Salmonella* and *E. coli*; Tn7 transposition was unsuccessful in these experiments.

Chromosomal integration of reporter fragments into the *att*Tn7 site downstream of the *glmS* gene of *S.* Typhimurium or *E. coli* was confirmed by PCR using the following primer sets: Reaction (1) glmSdetectFor + Lux-check, mCherry-check or GFP-check, Reaction (2) glmSdetectRev + Cm-check or Kn-check (Table 3.1). PCR products were purified using QIAquick PCR purification kits (Qiagen) and DNA sequencing was performed (Eurofins MWG Operon). To ensure the absence of secondary mutations in *Salmonella* reporter strains that were used for murine infection experiments, the reporter fragments (~3800 – 8800 bp) were moved from *S.* Typhimurium Δ*csgD* into the *S.* Typhimurium 14028 wild-type strain by P22 phage transduction (35). These P22 phage lysates are also available upon request.

3.4.7 Determination of *tnsABCD* transcript levels

RNA was isolated using an RNeasy Mini Kit (Qiagen) from *S.* Typhimurium 14028 Δ*csgD* cells transformed with pHSG415-*tnsABCD*, pTNS2 or pTNS3 and allowed to recover for 1h or 2h at 30ºC in SOC (as described in previous section). The quantity and purity of RNA was verified using a NanoDrop ND-1000 *s*pectrophotometer (Fisher Scientific) and integrity verified on a Agilent 2100 Bioanalyzer with a Prokaryote Total RNA Nano chip NanoDrop ND-1000 (Thermo Fisher Scientific). A total of 2ug RNA was then reverse transcribed using the High Capacity cDNA Reverse Transcription Kit (Thermo Fisher Scientific). Real time PCR reactions specific for *tnsA*, *tnsB*, *tnsC* or *tnsD* (primers listed in Table 3.1) were performed on 20ng equivalent cDNA using KAPA SYBR FAST qPCR mastermix (D-Mark Biosciences) on a Step-One-Plus Real-Time PCR system (Thermo Fisher Scientific). Analysis was conducted using the 2-ΔΔCT

method(36) with normalization against the geometric mean of two housekeeping genes, *fabG* and *groEL*.

3.4.8 Murine Infections with *S.* Typhimurium reporter strains.

Female C57BL/6 mice (6 to 8 weeks old) were purchased from Jackson Laboratory (Bar Harbor, ME) and housed in the VIDO animal facilities in direct accordance with guidelines drafted by the University of Saskatchewan's Animal Care Committee and the Canadian Council on the Use of Laboratory Animals. For competitive index (CI) experiments, two groups of six mice were assigned to groups using a randomization table prepared in Microsoft Excel; individual mice were marked with ear notches. Mice were infected by oral gavage with a mixed inoculum containing a 1:1 ratio of Kn^R and Cm^R *S.* Typhimurium 14028 strains containing the sig70_16-*luxCDABE* construct. An inoculum of 1.1×10^7 CFU, containing 5.3×10^6 Kn^R cells and 5.7×10^6 Cm^R cells, was delivered to each mouse in 100 µL of 100 mM HEPES pH 8. Infected mice were weighed daily and mice that had a >20% drop in weight, typically 4-7 days post-infection, were humanely euthanized. Spleen, liver, mesenteric lymph nodes (MLN), and cecum were collected from each mouse, placed in a 2mL Eppendorf Safe-Lock tube containing 1 mL of PBS and a 5mm steel bead (Qiagen product #69989) and were homogenized using a MM400 high-speed mixer mill (Retsch). For determination of total *Salmonella* CFU, organ homogenates were serially diluted in PBS and plated on Luria agar supplemented with 50 µg mL^{-1} Kn or 10 µg mL^{-1} Cm. The CI values were calculated as $[\text{CFU planktonic/CFU biofilm}]_{output}/[\text{CFU planktonic/CFU biofilm}]_{input}$.

For whole animal imaging experiments performed to compare light production between different reporter strains, groups of (n=4) C57BL/6 mice were pre-treated with an oral dose of 20 mg streptomycin 24 h prior to challenge (37) with 1×10^7 CFU of individual reporter strains: Cm^R sig70_16-, sig70c10- and sig70c35-*luxCDABE* or Kn^R-sig70c35-*luxCDABE*. Post-infection monitoring was performed as described above. Mice were imaged daily using the IVIS Lumina II (Perkin Elmer) and final tissue images were obtained after euthanization.

3.4.9 Complementation of *S.* Typhimurium 14028 Δ*csgD* mutant

For plasmid-based complementation, the region of DNA including the *csgD* ORF and intergenic region between the divergently transcribed *csgBAC* and *csgDEFG* operons was PCR amplified from *S.* Typhimurium 14028 genomic DNA using primers csgDcompFor2 and csgDcompREV. The resulting PCR product was purified and digested with *Eco*RI and *Ava*I and ligated into *Eco*RI/*Ava*I-digested pACYC184 (38) to generate pACYC-*csgD*. Product 1, consisting of *csgD* and corresponding intergenic region, was PCR-amplified from pACYC-*csgD* using primers csgDFOR_Bam and glmScsgDREV_Pst. Product 2, containing the chloramphenicol acetyltransferase (*cat*) gene and flanking Flp recombinase target (FRT) sites, was PCR-amplified from pKD3(39) using primers glmSCAMfor_Eco and CAMrev_Bam. Phusion High-Fidelity DNA polymerase was used for all PCR reactions (ThermoFisher Scientific). *Eco*RI/*Bam*HI-digested product 2 and *Bam*HI/*Pst*I-digested product 1 were sequentially ligated into *Eco*RI-*Bam*HI-*Pst*I-digested pTZ18U (Sigma-Aldrich). The *cat-csgD* product was PCR-amplified from pTZ18U using primers CAMfor_Eco and csgDREV_Kpn, digested with *Eco*RI and *Kpn*I and ligated into *Eco*RI/*Kpn*I-digested pUC18R6K-mini-Tn7T. The resulting vector was electroporated into *S.* Typhimurium Δ*csgD* cells containing the pHSG415-*tnsABCD* helper plasmid. Chromosomal insertion strains were selected by growth on Luria agar supplemented with 10 μg mL^{-1} Cm; final confirmation was performed by PCR using primers glmSdetectFor and glmSdetectRev followed by DNA sequencing.

3.4.10 Statistical Analysis

Maximum luciferase, mCherry or GFP expression values from *E. coli* or *S.* Typhimurium strains harboring pCS26 plasmids, or from *E. coli* APEC O1 or *S.* Typhimurium 14028 reporter strains, were compared between Kn^R or Cm^R, or sig70_16, sig70c10 or sig70c35 groups using unpaired t Tests with Welch's correction, which does not assume that groups have equal standard deviations. Competitive index values calculated from the organs of all 12 mice were first analyzed using a Shapiro-Wilk normality test, which determined that the data were not normally distributed. Wilcoxon

signed rank tests determined that the CI values for each organ were not statistically different (p >0.05) from a theoretical value of 1.0. All analyses were performed using GraphPad Prism v6.0.

3.5 Results

3.5.1 Modular pCS26 vectors with synthetic, σ^{70}-dependent promoters

The pCS26 vector, originally developed in the laboratory of Dr. Michael Surette (22) (http://www.surettelab.ca/lab/), has been used for several large-scale gene expression experiments (4, 22-24). pCS26 offers flexibility in both the promoters and reporters that can be incorporated (Fig. 3.1A). We generated versions of pCS26 containing either mCherry or GFP reporters, in addition to the base plasmid containing the *luxCDABE* operon from *Photorhabdus luminescens.* To increase the utility of the pCS26 reporter system and enable experiments where strains can be combined, such as competitive infections or fitness assays, we generated a version of pCS26 with chloramphenicol-resistance as a selectable marker (see Materials and Methods). We incorporated promoters of increasing strength based on a constitutive, σ^{70}-dependent promoter named sig70_16 (40). This promoter was designed based on the consensus sequence of different σ^{70}–controlled promoters in *E. coli,* with degenerate positions in the -10 and -35 regions to introduce variability in promoter strength (K. Pabbaraju and MG Surette, unpublished data). We reasoned that restoring the sig70_16 sequence back to the consensus -10 or -35 sequence would boost the promoter strength. We named these new promoters sig70c10, with the -10 sequence restored to consensus, and sig70c35, with the -35 sequence restored to consensus (Fig. 3.1A). Luciferase expression from these plasmids was monitored in *E. coli* DH10B and *S.* Typhimurium 14028 during a 48-hour growth period (Fig. 3.2). In general, the profiles differentiated into three distinct groups with the sig70-16 promoter having the lowest expression, sig70c10 with intermediate levels and sig70c35 with the highest expression. There was very little difference in expression between the Kn^R and Cm^R plasmids, except in *Salmonella* where the Cm^R sig70c35 plasmid construct had a lower maximum expression level (Fig. 3.2).

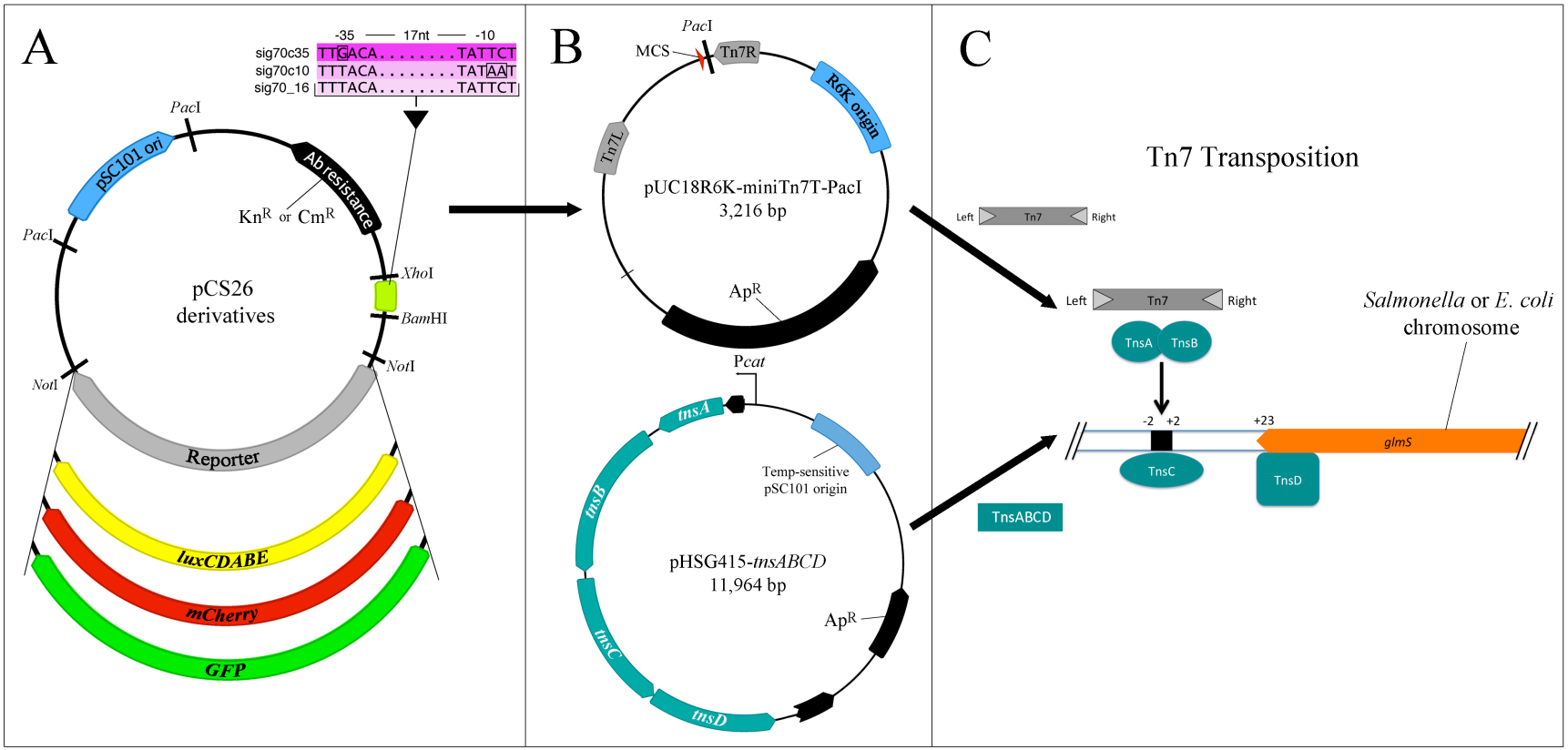
A
-35
17nt
-10
sig70c35 TTGACA.....TATTCT
sig70c10 TTTACA.....TATAAT
sig70_16 TTTACA.....TATTCT
PacI
pSC101 ori
Ab resistance
KnR or CmR
PacI
XhoI
BamHI
pCS26
derivatives
NotI
NotI
Reporter
luxCDABE
mCherry
GFP
B
PacI
MCS
Tn7R
R6K origin
Tn7L
pUC18R6K-miniTn7T-PacI
3,216 bp
ApR
Pcat
tnsA
tnsB
Temp-sensitive
pSC101 origin
pHSG415-tnsABCD
11,964 bp
ApR
tnsC
tnsD
C
Tn7 Transposition
Left
Tn7
Right
Left
Tn7
Right
TnsA
TnsB
-2
+2
+23
glmS
TnsC
TnsD
TnsABCD
Salmonella or E. coli
chromosome

Figure 3.1. Use of modular pCS26 plasmids in a modified Tn7 transposition system.

(A) Choose the antibiotic resistance marker (Kn^R or Cm^R), sig70 promoter (or new promoter cloned with *Xho*I/*Bam*HI) and *luxCDABE*, mCherry, or GFP (or new reporter cloned with *Not*I/*Not*I) in the modular pCS26 plasmid. The synthetic, σ^{70}-dependent promoters (shown in magenta) allow for constitutive expression at differing strengths; the nucleotide differences are outlined, and -35 and -10 regions are recognized by σ^{70}. (B) pCS26 constructs are digested with *Pac*I and ligated into the pUC18R6K-miniTn7T-PacI delivery plasmid. pHSG415-*tnsABCD* is a new helper plasmid that is replication proficient and can be conditionally maintained in most Enterobacteriaceae strains. The *tnsABCD* operon is expressed under control of the pre-existing *cat* promoter, leading to elevated expression; two black arrows represent truncated *aph* (Kn^R) and *cat* (Cm^R) genes. (C) 100 ng of miniTn7T-pCS26 delivery plasmid is transformed into cells containing pHSG415-*tnsABCD*, resulting in orientation-specific integration of Kn^R or Cm^R-promoter-reporter constructs into the *att*Tn7 site (shown as a black box) present in the chromosome of *Salmonella* and *E. coli*, downstream of the *glmS* gene. The details of Tn7 transposition were adapted from (16). Chromosomal insertion can be verified by PCR, followed by growth at 37ºC to cure the pHSG415 helper plasmid. All plasmid maps were generated using Geneious v. 9.0.4 (www.biomatters.com).

3.5.2 Modifying the Tn7 transposition system for chromosomal integration of luxCDABE constructs.

Plasmid stability is a significant issue when it comes to *in vivo* studies of microbial pathogenesis within a host. To expand the spectrum of uses for the modular, pCS26 reporter system, we wanted to develop a chromosomal integration strategy based on Tn7 transposition. The original Tn7 system, first described by Choi *et al.* (12) consists of two replication-sensitive vectors that are used to carry: (1) the gene(s) of interest and (2) the transposase machinery required for chromosomal integration. The first modification we made to the existing Tn7 system was to insert a *Pac*I polylinker into the multiple cloning site of the pUC18R6K miniTn7T delivery plasmid (Fig. 3.1B). This enabled cloning of the promoter, reporter and antibiotic resistance cassettes from pCS26-Kn^R or -Cm^R. We attempted to integrate the pCS26-*luxCDABE* constructs into the chromosome of *Salmonella* using one of the original helper plasmids, pTNS2(12) or pTNS3 (34), but the experiments were unsuccessful (Table 3.2). Since pTNS2 and pTNS3 do not replicate in *Salmonella*, we reasoned that *tnsABCD* expression might not reach high levels during the transformation process; this was subsequently confirmed by qRT- PCR (Table 3.2). Therefore, the second modification we made was to generate pHSG415-*tnsABCD*, a replication-proficient helper plasmid that contains a temperature-sensitive origin of replication to allow for curing of cells. Use of this plasmid, with *tnsABCD* genes cloned under partial control of the chloramphenicol promoter (Fig. 3.1B), resulted in 10-30 fold enhanced gene expression as measured by qRT-PCR (Table 3.2). We achieved Tn7 transposition of the sig70-*luxCDABE* reporter constructs into the genome *S.* Typhimurium 14028 (Fig. 3.1C) at high efficiency (Table 3.2). Since ligation of *Pac*I-digested pCS26 reporter constructs into the miniTn7T delivery vector resulted in two possible insert orientations, we used the orientation that had slightly stronger expression (Fig. 3.3). For *E. coli* APEC O1, transposition with pHSG415-*tnsABCD* was even more efficient than in *Salmonella*, ranging from 2.2×10^{-8} to 1.4×10^{-10}. In contrast, the use of pTNS2 or pTNS3 did not result in any successful transposition events.

We analyzed the *S.* Typhimurium 14028 and *E. coli* APEC O1 *luxCDABE* reporter strains during growth in liquid culture at 37ºC. Expression profiles clustered into

Table 3.2. *tnsABCD* expression from different Tn7 helper plasmids and resulting transposition frequencies in *S.* Typhimurium 14028.

	***tnsABCD* expression[b]**		**Transposition frequency[c]**		
Helper plasmid[a]	**1 h**	**2 h**	***luxCDABE***	**mCherry**	**GFP**
pTNS2	1.0 ±0.1	1.9 ±0.5	0	0	0
pTNS3	2.4 ±1.0	4.2 ±1.2	0	0	0
pHSG415-*tnsABCD*	31.9 ±3.8	33.5 ±3.6	6.0×10^{-9} to 3.8×10^{-10}	5.6×10^{-9} to 3.8×10^{-10}	6.0×10^{-9} to 7.5×10^{-10}

[a] 100ng of pTNS2 or pTNS3 were electroporated into *S.* Typhimurium 14028 competent cells. For pHSG415-*tnsABCD*, competent cells were prepared containing this helper plasmid.

[b] Total RNA was isolated from each competent cell / helper plasmid combination after recovery in SOC medium at 30ºC for 1 h or 2 h. Expression levels were calculated individually for *tnsA,B,C* and *D* as described in the Materials and Methods; values represent the average fold-change for all four genes measured from four biological replicate samples, plus or minus the standard deviation.

[c] For each reporter type, 100ng of pUC18R6K-mini-Tn7T-pCS26 delivery plasmid, containing one of six different constructs (Kn^R or Cm^R with sig70_16, sig70c10 or sig70c35 promoters), was electroporated into *S.* Typhimurium 14028 cells containing the different helper plasmids. The values represent the range of transposition frequencies calculated based on the number of luminescent or fluorescent Kn^R or Cm^R colonies obtained from a total number of 2.7×10^9 competent cells. No colonies were obtained using pTNS2 or pTNS3 helper plasmids.

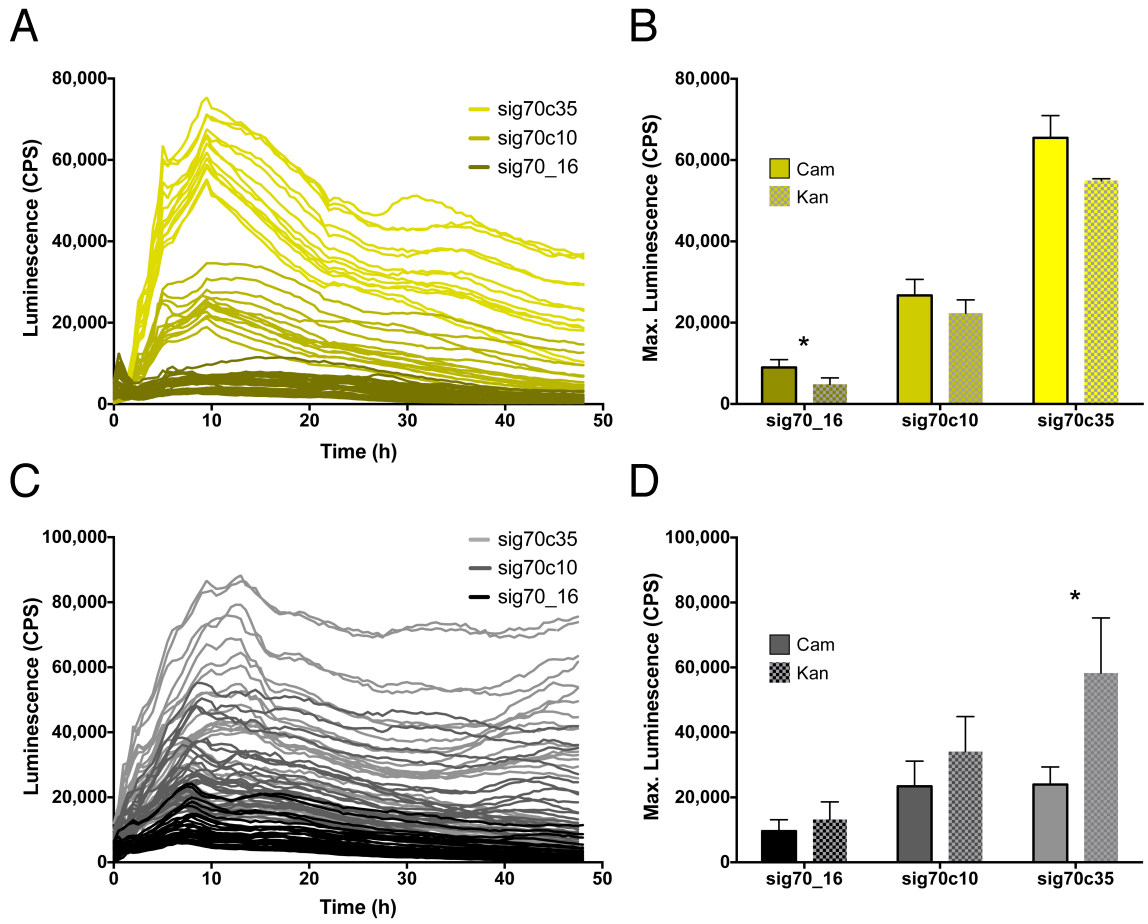

Figure 3.2. Plasmid-based luciferase expression in *E. coli* and *S.* Typhimurium. pCS26 plasmids expressing luxCDABE under the control of synthetic, σ^{70}-dependent promoters were transformed into *E. coli* DH10B (A) or S. Typhimurium 14028 (C) and luciferase expression (counts per second; CPS) was measured during growth in LB at 37ºC for 48 h. (B and D) Bars show the mean and standard deviation of the maximum expression values corresponding to the CmR and KnR plasmids within each sig70 promoter group. Statistical significance (*) between CmR and KnR groups was determined using multiple t tests, corrected for multiple comparisons using the Holm-Sidak method (alpha = 5%). The CmR-sig70c35 pCS26 plasmid had low expression in S. Typhimurium even when strains were grown in the presence of less chloramphenicol (i.e., 5, 7, or 9 µg mL-1; data not shown).

three distinct groups based on promoter strength (Fig. 3.4A, D). For *S.* Typhimurium 14028, the peak expression values were 8,000 counts per second (CPS) for sig70_16, 30,000 CPS for sig70c10, and 62,000 CPS for sig70c35, which represented step-wise increases of 3.6x and 8.2x (Fig. 3.4B). For *E. coli* APEC O1, the increase between sig70_16 and sig70c10 was 2.4x and between sig70c10 and sig70c35 was 5.6x (Fig. 3.4E). For both *S.* Typhimurium and *E. coli*, there were only small expression differences between the Kn^R and Cm^R constructs (Fig. 3.4C, F). These results confirmed that the modularity of the pCS26 vector system was conserved upon integration into the chromosome.

3.5.3 Generating fluorescent *S.* Typhimurium and *E. coli* reporter strains.

Fluorescent reporters can be used to visualize bacterial cells by microscopy after tissue sectioning or in *in vitro* experiments. To this end, we generated mCherry- and GFP-expressing *S.* Typhimurium reporter strains at high efficiency (Table 3.2). Expression from these strains was measured after 24 h growth at 37 ºC on Luria agar (Fig. 3.5A). Increasing levels of expression with the sig70_16, sig70c10 and sig70c35 promoters could be measured for each reporter (Fig. 3.5B). The dynamic range of GFP and mCherry was reduced compared to luciferase, however, a more observable difference between promoters was measured during growth in liquid culture (Fig. 3.6). Importantly, only minor background fluorescence for GFP-expressing strains was detected in the mCherry channel and vice versa (Fig. 3.5A). A confocal microscopy image was generated for the Kn^R *S.* Typhimurium sig70_16-GFP reporter strain, which demonstrated that individual cells within the population were expressing similar levels of GFP (Fig. 5.3C). Similar images were generated for the *S.* Typhimurium mCherry reporter strains (data not shown).

mCherry- and GFP-reporter strains were also generated for *E. coli* APEC O1 at efficiencies as high as 2.2×10^{-8} (data not shown). As in *S.* Typhimurium, the expression profiles of the σ^{70}-dependent promoters were distributed into three distinct groups during growth in liquid culture (Fig. 3.7). Fluorescent APEC O1 cells were also visualized by confocal microscopy (Fig. 3.7).

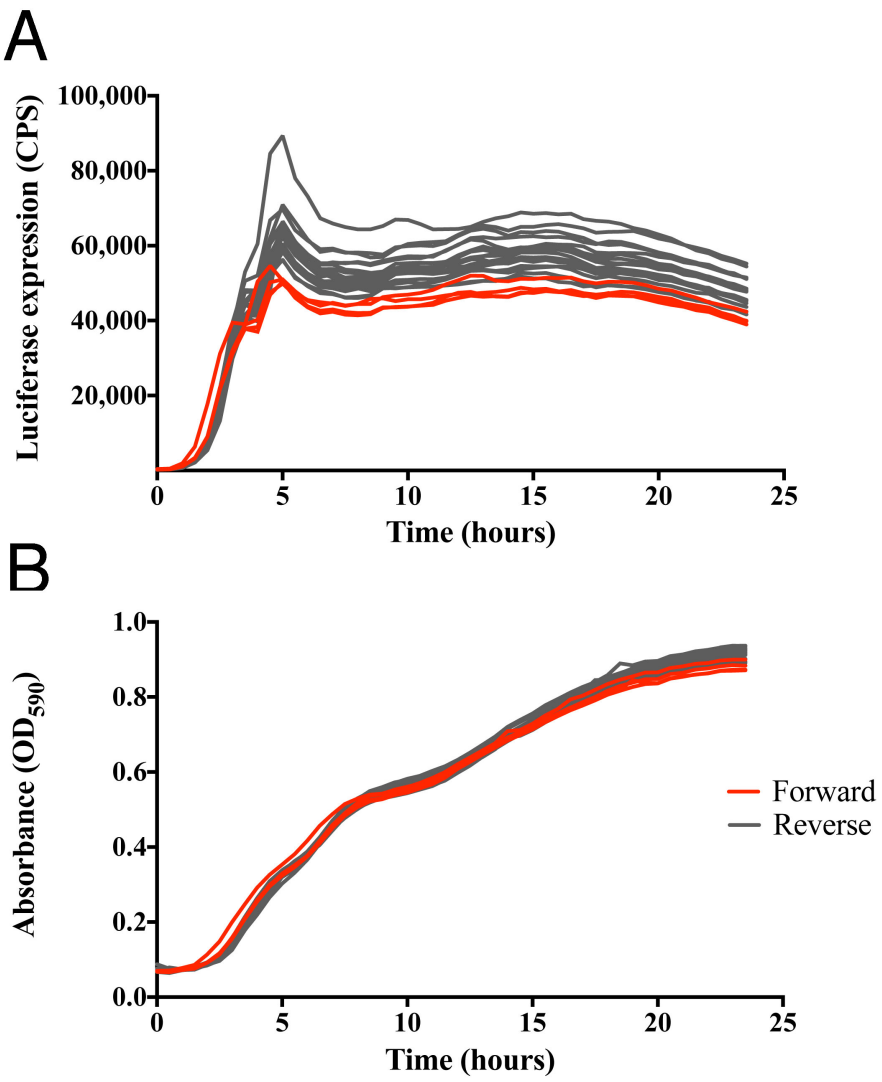

Figure 3.3. The influence of construct orientation on luciferase expression in *S.* Typhimurium.

Ligation of *Pac*I-digested pCS26 reporter constructs into the pUC18R6K miniTn7T-PacI delivery vector resulted in two possible insert orientations: what we termed forward (Tn7L-Kn^R/Cm^R-promoter-*luxCDABE*-Tn7R) and reverse (Tn7L-*luxCDABE*-promoter-Kn^R/Cm^R-Tn7R). Since Tn7 transposition is orientation-specific, this lead to the generation of two final strain possibilities. We wanted to determine if construct orientation in the *Salmonella* chromosome affected *luxCDABE* expression. (A) Luciferase expression (counts per second; CPS) and (B) optical density at 590 nm were measured every 30 min during growth of *S.* Typhimurium Kn^R or Cm^R reporter strains

containing the sig70_16-*luxCDABE* construct integrated in the forward (Red) or reverse (Grey) orientations. Each line represents the mean luciferase expression at each time point from one individual clone grown in triplicate. Four forward orientation clones and 18 reverse orientation clones were assayed.

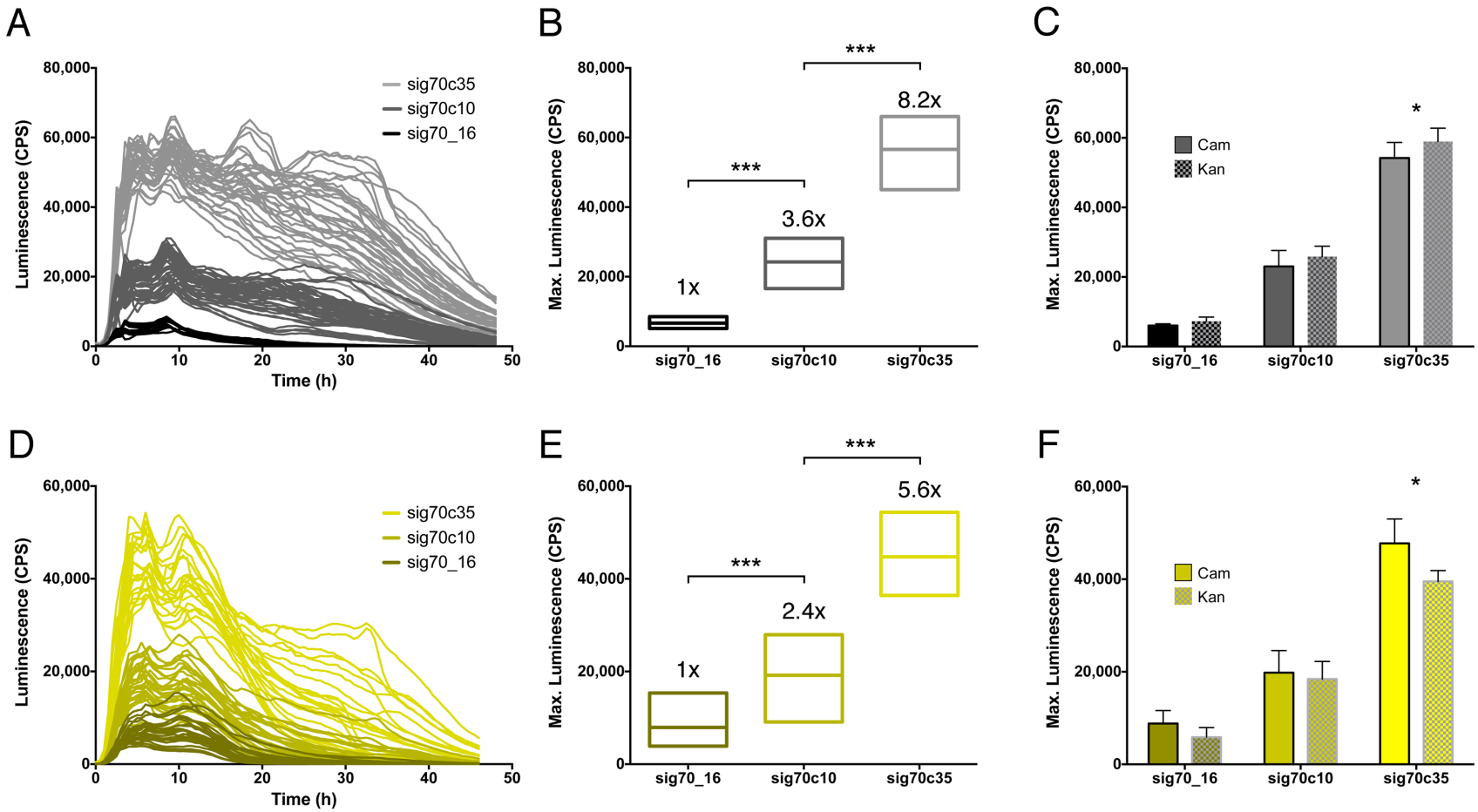

Figure 3.4. Chromosome-based expression of luciferase in *S.* Typhimurium and *E. coli*.

Synthetic, σ^{70}-dependent promoter - *luxCDABE* operon fusions were integrated into the chromosomes of *S.* Typhimurium 14028 Δ*csgD* or *E. coli* APEC O1 using the modified Tn7 transposition system. (A and D) Luciferase expression from each strain was

measured during growth at 37ºC for 48 h; each line represents the expression curve from one individual reporter strain. (B and E) Box plots show the mean and range of maximum expression values for the different chromosomal reporter strains. Numerical values above the boxes represent the step-wise increase in mean expression levels between promoters. Statistical significance was calculated using unpaired *t* tests with Welch's correction (***, $p < 0.001$). (C and F) Maximum expression values for Cm^R and Kn^R strains within each sig70 promoter group are shown with histogram bars representing the mean and error bars representing the standard deviation. Statistical significance (*) was determined using multiple *t* tests, corrected for multiple comparisons using the Holm-Sidak method (alpha=5%).

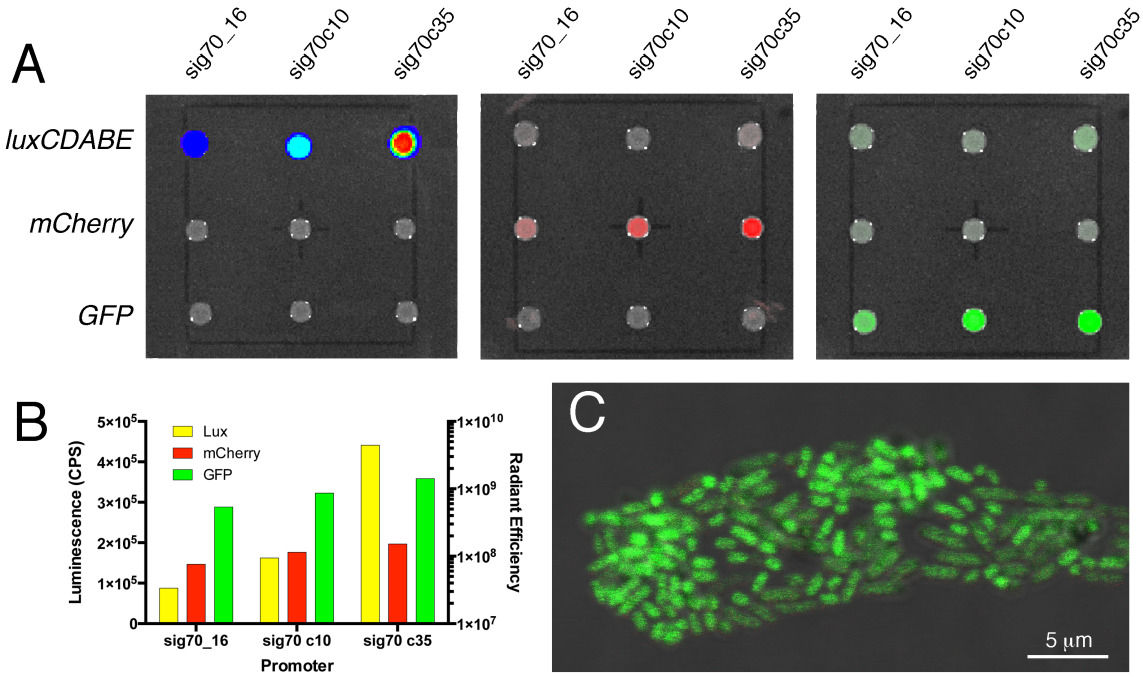

Figure 3.5. Fluorescent reporter strains of *Salmonella* serovar Typhimurium 14028. (A) *S.* Typhimurium reporter strains with sig70_16, sig70c10, or sig70c35 promoters controlling *luxCDABE*, mCherry or GFP expression were grown on Luria agar at 37°C. Expression from the different strains was visualized using a IVIS Lumina II whole animal imaging system (Perkin-Elmer). B) Peak luciferase (CPS) or fluorescence (radiant efficiency) was measured using the region of interest (ROI) module in Living Image software ver. 4.2 (Perkin Elmer). C) Cells of the Kn^R sig70_16 GFP reporter strain were visualized on a Leica TCS SP5 Confocal microscope, using the 63x oil immersion objective with the 180 laser operating at 488 nm. The scale bar is displayed in the bottom right corner.

3.5.4 Use of luciferase-expressing *S.* Typhimurium reporter strains *in vivo*.

S. Typhimurium is a well-established pathogen capable of causing disease in susceptible mouse strains (41). We wanted to assess the virulence of the bacterial luciferase reporter strains, and determine if the antibiotic resistance markers would have any influence. To ensure that Kn^R and Cm^R strains would be recovered with equal efficiencies, we tested their plating efficiencies on Luria agar supplemented with 7, 10 or 15 μg mL^{-1} of Cm or 50 or 100 μg mL^{-1} of Kn. Similar CFU values were detected for the strains on all media used, indicating that the antibiotic concentration did not affect their recovery (Table 3.3). C57BL/6 mice were challenged with a 1:1 ratio of Kn^R and Cm^R sig70_16-*luxCDABE* reporter strains. At 4-7 days post-infection, infected mice were euthanized and the bacterial loads were determined for the spleen, liver, cecum and mesenteric lymph nodes (Fig. 3.8). In some individual mice, the Kn^R reporter strain (black bars) was recovered at higher levels while in other mice the Cm^R reporter strain (grey bars) was recovered at higher levels. Typically, the same strain was recovered at higher levels in all organs from a single mouse, suggesting that the outcome of the competition between the strains was decided early on during infection. The competitive index (CI) values were calculated from each individual mouse; for each organ, the values were not significantly different than 1 (Fig. 3.8E). Based on this experiment, we concluded that the Kn^R and Cm^R *S.* Typhimurium reporter strains were equally virulent in this model.

In a separate trial, C57BL/6 mice were orally infected with individual *S.* Typhimurium sig70_16, sig70c10 and sig70c35 luciferase reporter strains and visualized using a whole animal imager. Each reporter strain produced a luciferase signal that was detectable within whole mice at day 1 post-infection and in their internal organs after 3-6 days post-infection (Fig. 3.9). Similar to the *in vitro* experiments, there were step-wise increases in expression between the three promoters and the sig70c35 reporter strain had the highest levels of light production (Fig. 3.9C, F). Luciferase production was also detected from *S.* Typhimurium cells shed into the fecal pellets (data not shown).

Table 3.3. Plating efficiency of *S.* Typhimurium reporter strains on various media.

	Number of CFU obtained on each media type[a]					
Strain[b]	**No Ab**	**Cm_7**	**Cm_{10}**	**Cm_{15}**	**Kn_{50}**	**Kn_{100}**
WT Kn^R	3.75×10^8	0	0	0	3.88×10^8	3.21×10^8
WT Cm^R	4.42×10^8	4.17×10^8	4.75×10^8	4.79×10^8	0	0
WT Kn^R/WT Cm^R	3.83×10^8	2.42×10^8	2.42×10^8	2.33×10^8	1.71×10^8	1.71×10^8

[a] Inocula were planted on various media and the recovery of each strain was calculated. The values represent the mean CFU calculated from each inoculum; serial dilutions and plating performed using the drop dilution method. Media consisted of Luria agar with no antibiotic (No Ab) or supplemented with 7, 10 or 15 μg mL^{-1} chloramphenicol or 50 or 100 μg mL^{-1} kanamycin.

[b] *S.* Typhimurium 14028 reporter strains contained the Kn^R or Cm^R sig70_16-*luxCDABE* reporter construct integrated into the genome. Inocula consisted either of an individual strain or the strains combined at a 1:1 ratio.

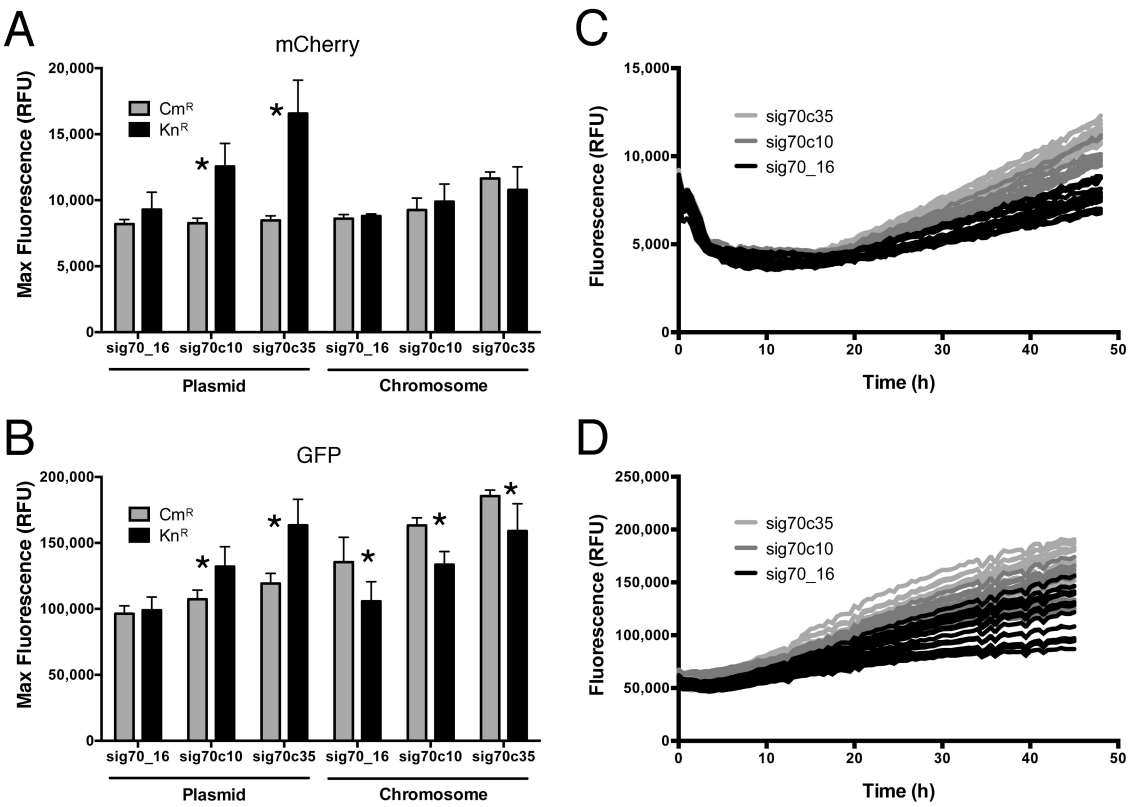

Figure 3.6. The influence of antibiotic resistance marker and promoter strength on mCherry and GFP expression in *S.* Typhimurium.

Peak fluorescence (relative fluorescent units, RFU) was measured from pCS26 plasmids containing the sig70-16, sig70c10 or sig70c35 promoters and mCherry (A) or GFP (B) reporters in *S.* Typhimurium 14028 or from constructs that were integrated into the chromosome of *S.* Typhimurium 14028 Δ*csgD*. Bars represent the mean and standard deviation from at least three biological replicates. Statistical significance (*) between Cm^R and Kn^R strains was determined using multiple *t* tests, corrected for multiple comparisons using the Holm-Sidak method (alpha of 5%). (C and D). 48 h expression profiles for *S.* Typhimurium reporter strains with mCherry or GFP constructs integrated into the chromosome. Each line represents the mean luciferase expression at each time point from one individual clone grown in triplicate. All strains were grown in LB at 37° C for 48 h.

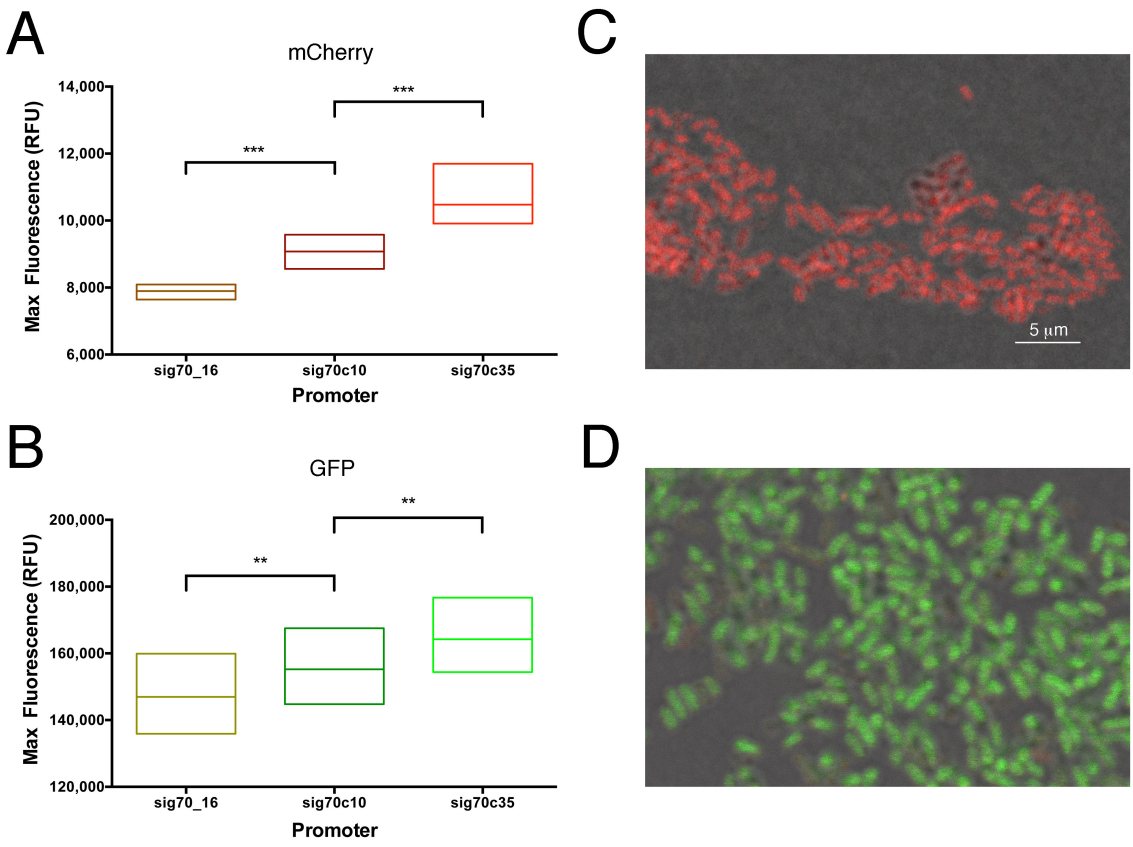

Figure 3.7. Chromosomal mCherry and GFP reporters in *E. coli* APEC O1.

Peak fluorescence (relative fluorescent units, RFU) was measured from *E. coli* APEC O1 reporter strains containing sig70-16, sig70c10 or sig70c35 promoters and mCherry (A) or GFP (B) reporters integrated into the genome. Box plots show the range of expression from multiple biological replicates with the mean represented by the horizontal line within each box. Statistical significance between each promoter group was determined using unpaired *t* tests with Welch's correction (**, p <0.01; ***, p <0.001). (C and D) Cells of the sig70c10-mCherry or sig70c10-GFP Cm^R reporter strains were visualized on a Leica TCS SP5 Confocal microscope, using the 63x oil immersion objective with the 180 laser operating at 488 nm. The scale bar is displayed in the bottom right corner of (C).

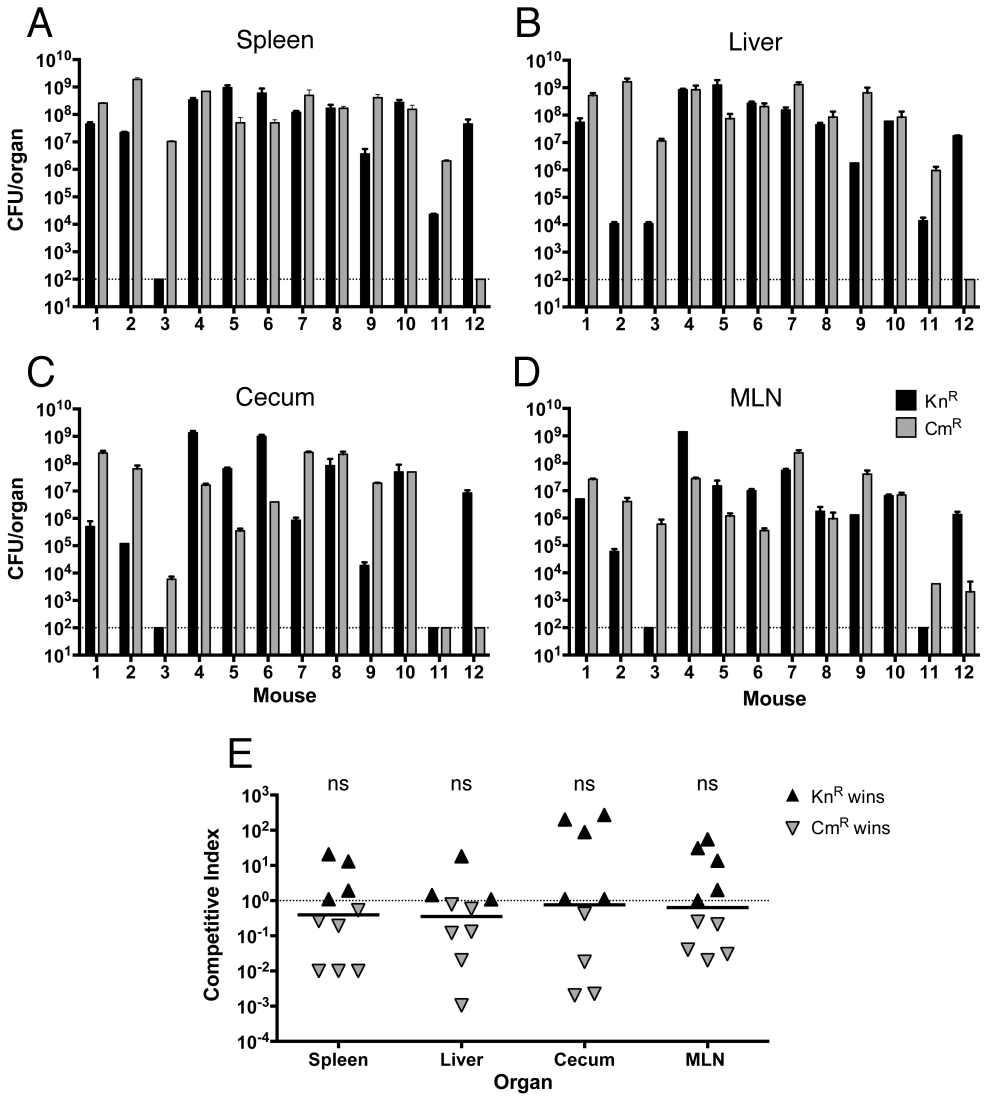

Figure 3.8. Bacterial counts and competitive indices from C57BL/6 mice after oral infection with *S.* Typhimurium sig70_16 *lux* operon fusion strains.

The spleen (A), liver (B), cecum (C), and MLN (D) were collected from euthanized mice, homogenized and plated on Luria agar supplemented with 50 µg mL^{-1} Kn or 10 µg mL^{-1} Cm to determine the levels of each strain present (measured as CFU per whole organ). E) Competitive index (CI) values were calculated for each mouse in each organ: CI= ([CFU Kn^R OUT/CFU Cm^R OUT]/[CFU Kn^RIN/CFU Cm^RIN]). A CI value of 1.0, which indicates equal virulence, is represented by a horizontal dotted line. Statistical significance from a value of 1 was determined using a Wilcoxon Signed Rank Test (ns; not significant, $p > 0.05$).

3.5.5 Chromosomal complementation of a *S.* Typhimurium Δ*csgD* mutant strain.

We predicted that we could use the modified Tn7 system to complement a *S.* Typhimurium Δ*csgD* strain at near stoichiometric expression levels. Regulation of *csgD* expression is known to be complex, involving multiple transcriptional regulatory proteins, sRNAs and histone-like proteins, such as H-NS and IHF (42). With *csgD* and its native promoter inserted downstream of *glmS* in the chromosome, the resulting *csgD*$^+$ strain was compared to the Δ*csgD* strain complemented from a plasmid (i.e., pACYC-*csgD*).

CsgD is the central regulator of biofilm development in *Salmonella* and it controls several multicellular behaviors that have been collectively termed the red, dry and rough (rdar) morphotype (43). Rdar morphotype colonies are adhesive, patterned colonies that can be lifted intact off the agar surface after 48h of growth (26, 43). As expected, the WT *S.* Typhimurium 14028 strain formed rdar colonies, whereas the Δ*csgD* mutant strain formed mucoid colonies (Fig. 3.10A). Colony morphology of the *csgD*$^+$ strain closely matched the WT strain (Fig. 3.10A), which was consistent with production of curli fimbriae and cellulose, two of the extracellular matrix components that are regulated by CsgD (44). The pACYC-*csgD* strain yielded colonies that were larger and had a flatter appearance at 30ºC (Fig. 3.10A) and were rdar-like at 37ºC, which was consistent with over-production of curli and cellulose (43). We also observed the phenotypes of each strain grown in liquid culture under biofilm-inducing conditions; cells are known to differentiate into multicellular aggregates (CsgD-ON state) and planktonic cells (CsgD-OFF state) (45, 46). Multicellular aggregates were formed in WT cultures and were notably absent in the Δ*csgD* strain (Fig. 3.10B). Aggregates were visible in the *csgD*+ and pACYC-*csgD* strains, demonstrating that CsgD expression had been restored. However, the pACYC-*csgD* cultures appeared almost entirely devoid of planktonic cells (Fig. 3.11).

To quantitate the degree of CsgD complementation in each strain, we measured the expression of *csgD* and *csgB*, coding for curli fimbriae production, and *adrA*, coding for a regulator of cellulose production, using promoter-*lux* operon fusion plasmids. In addition, the activity level of RpoS, a sigma factor that is known to regulate *csgD*

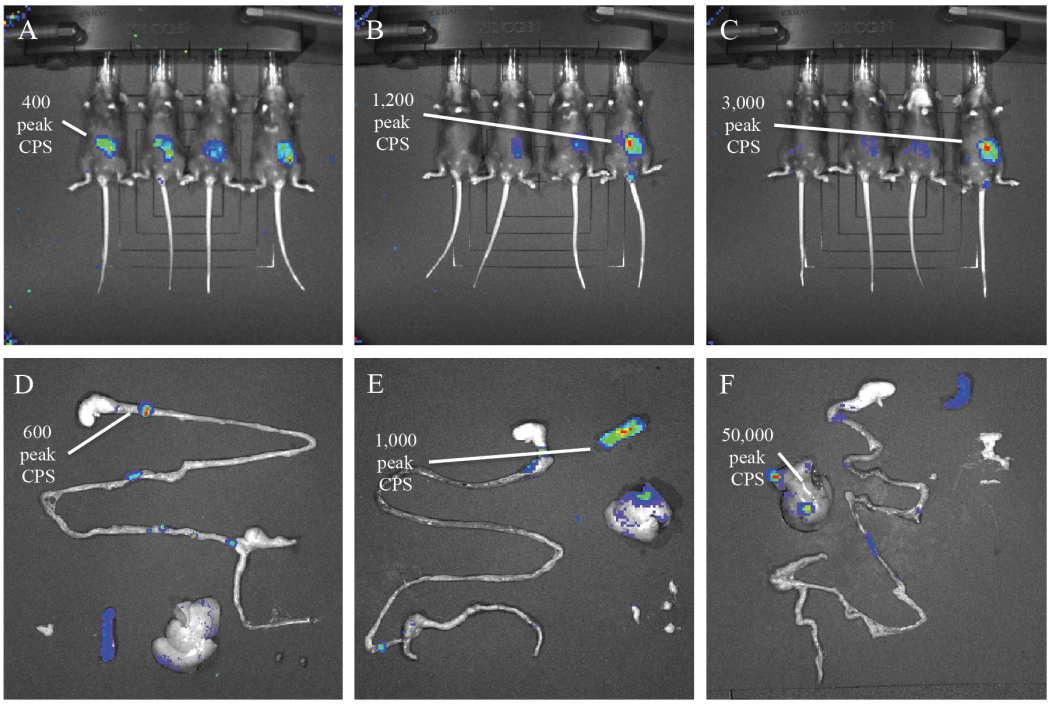

Figure 3.9. *S.* Typhimurium reporter strains have varying levels of luciferase expression *in vivo*.

C57BL/6 mice were pre-treated with streptomycin then challenged with *S.* Typhimurium *luxCDABE* reporter strains containing promoters sig70_16 (A,D), sig70c10 (B,E) or sig70c35 (C,F). Mice were visualized on a whole animal imager at one day post-infection (A, B, C). At 4-5 days post-infection, mice were euthanized and their organs (spleen, liver, GI tract (stomach, intestine, cecum, and colon) and MLN) were recovered and imaged (D, E, F). Peak luciferase values are indicated (measured in CPS).

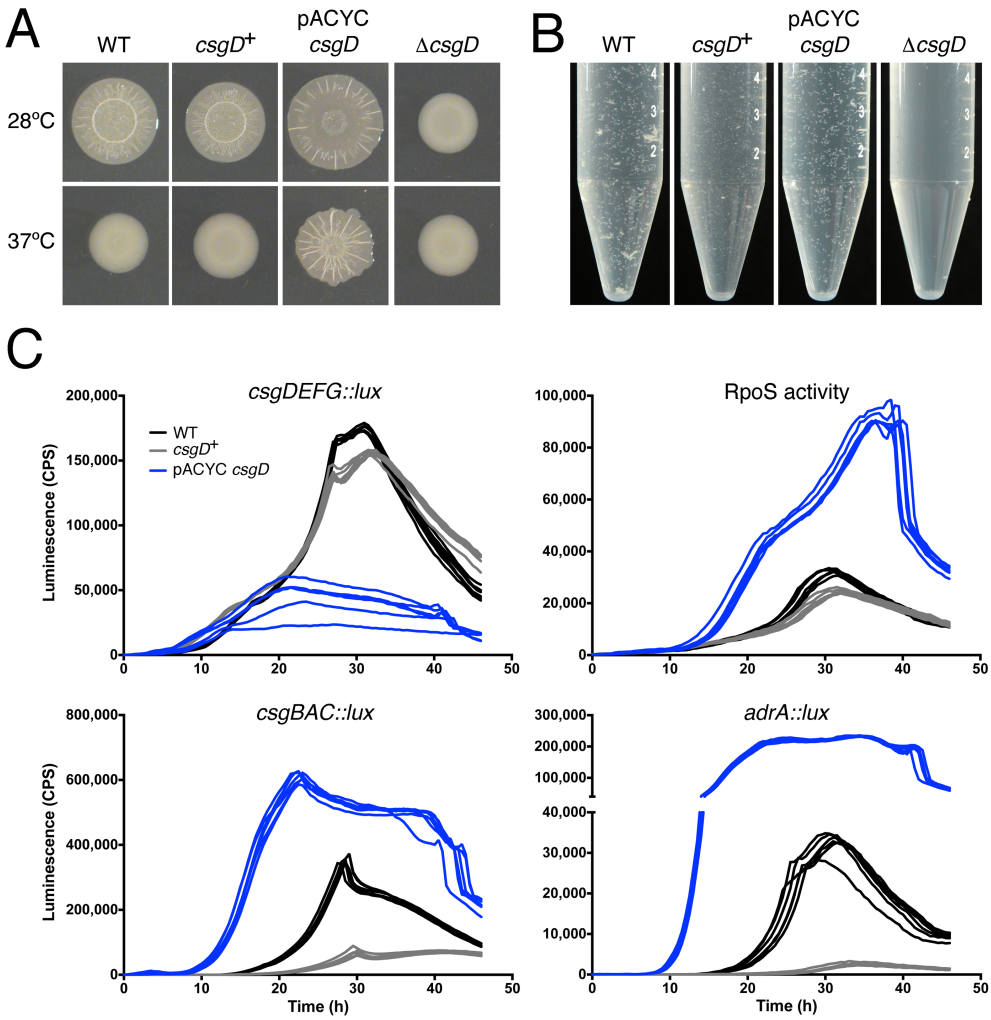

Figure 3.10. Chromosome-based complementation of a *S.* Typhimurium Δ*csgD* mutant strain.

S. Typhimurium ATCC 14028 (WT), Δ*csgD*, *csgD*+, and pACYC-*csgD* were grown on 1% tryptone agar at 28ºC or 37ºC (A) or in 1% tryptone broth at 28ºC (B). (C) *csgDEFG*, *csgBAC*, and *adrA* expression was measured during growth using promoter-luciferase (*luxCDABE*) fusion plasmids, designed to measure gene expression by light production (counts per second, CPS). RpoS activity was measured using the sig38H4::*lux* reporter plasmid that contains a synthetic, RpoS-dependent promoter (26). Each line represents the expression curve from one biological replicate culture performed in triplicate, where CPS was measured at half hour intervals during growth. Cells were grown in 1% tryptone broth at 28ºC with agitation; expression was measured in a Victor X^3 multi-label plate reader (Perkin-Elmer).

expression and several downstream components, was monitored using a synthetic, RpoS-responsive promoter (26, 28). RNA-seq transcriptome analysis indicated that the presence of a *luxCDABE* fusion plasmid only has a minor effect on cell physiology (Fig. 3.12), therefore, these plasmids provide a good measure of real-time gene expression. Expression of *csgD* and the RpoS-responsive promoter matched closely between the WT and *csgD*$^{+}$ strains during growth at 28ºC (Fig. 3.10C). In contrast, the plasmid-complemented strain had a different *csgD* expression profile and the RpoS activity was approximately three times higher (Fig. 3.10C, blue lines). Expression of *csgB* and *adrA* was similarly elevated in the pACYC-*csgD* strain (i.e., 2x higher for *csgB* and 10x higher for *adrA*) as compared to the WT. In the *csgD*$^{+}$ strain, *csgB* and *adrA* were expressed at lower levels, but the temporal profiles were similar to the WT strain (Fig. 3.10C). From this collection of experiments, we concluded that the chromosomally complemented strain had more stoichiometric CsgD expression than the plasmid-complemented strain.

3.6 Discussion

We developed a more efficient Tn7 transposition system for use in *Salmonella*, *E. coli* and other Enterobacterial strains. Transposition was successful in 40 separate experiments when using our pHSG415-*tnsABCD* helper plasmid. In comparison to the pre-existing Tn7 system, we did not achieve a single transposition event when using the non-replicating, helper plasmids pTNS2 (12) or pTNS3 (34). While these plasmids may still work if transformed via mating (47, 48), transposition is undoubtedly a rare event. Use of our helper plasmid resulted in 10-30x boosted *tnsABCD* expression compared to pTNS2 or pTNS3 and transposition frequencies ranging from 10^{-8} to 10^{-10}. pHSG415 is particularly useful because it is unstable in the absence of selection pressure (49), therefore growth at 37ºC without antibiotic should result in loss of plasmid from 100% of cells. We also tested a helper and delivery combination plasmid that was specifically designed for use in *Salmonella* (i.e., pGRG25) (18). Presumably due to the large 12.5 kb size of this vector, we found it difficult to clone into and we also did not achieve a single transposition event after 20 attempts. The only problem we encountered with pHSG415-*tnsABCD* was with an antibiotic-resistant *E. coli* poultry isolate that we were unable to transform (A. Dar and B. Allan, unpublished data). Strain resistance could be overcome

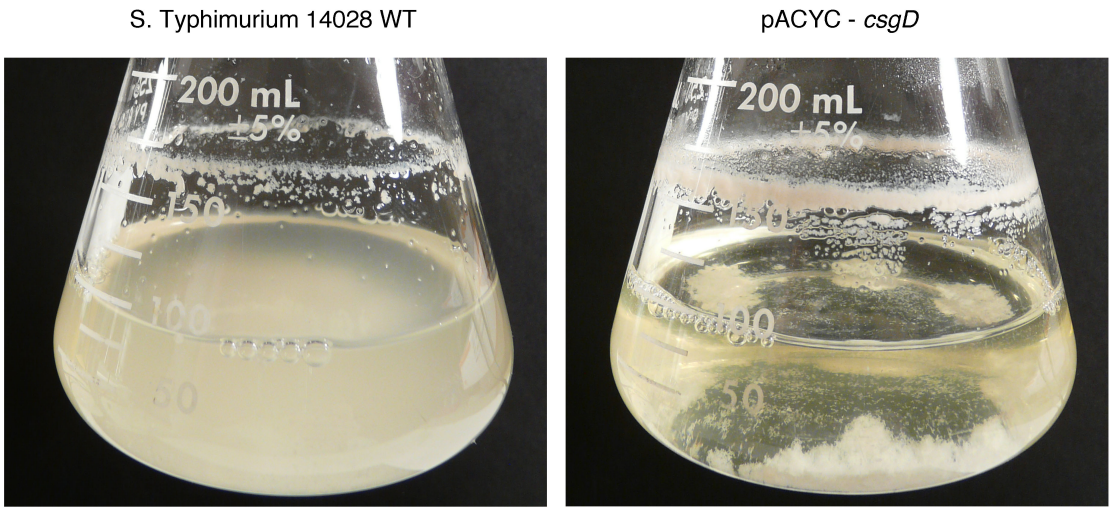

Figure 3.11. Comparison of wildtype *S.* Typhimurium 14028 and a plasmid-complemented Δ*csgD* strain in an *in vitro* flask model of biofilm development.
Strains were grown for 18 h in 1% tryptone broth at 28ºC with agitation. The media in the WT culture appears cloudy and contains the planktonic cell fraction, whereas the media in the pACYC-*csgD* strain is mostly clear.

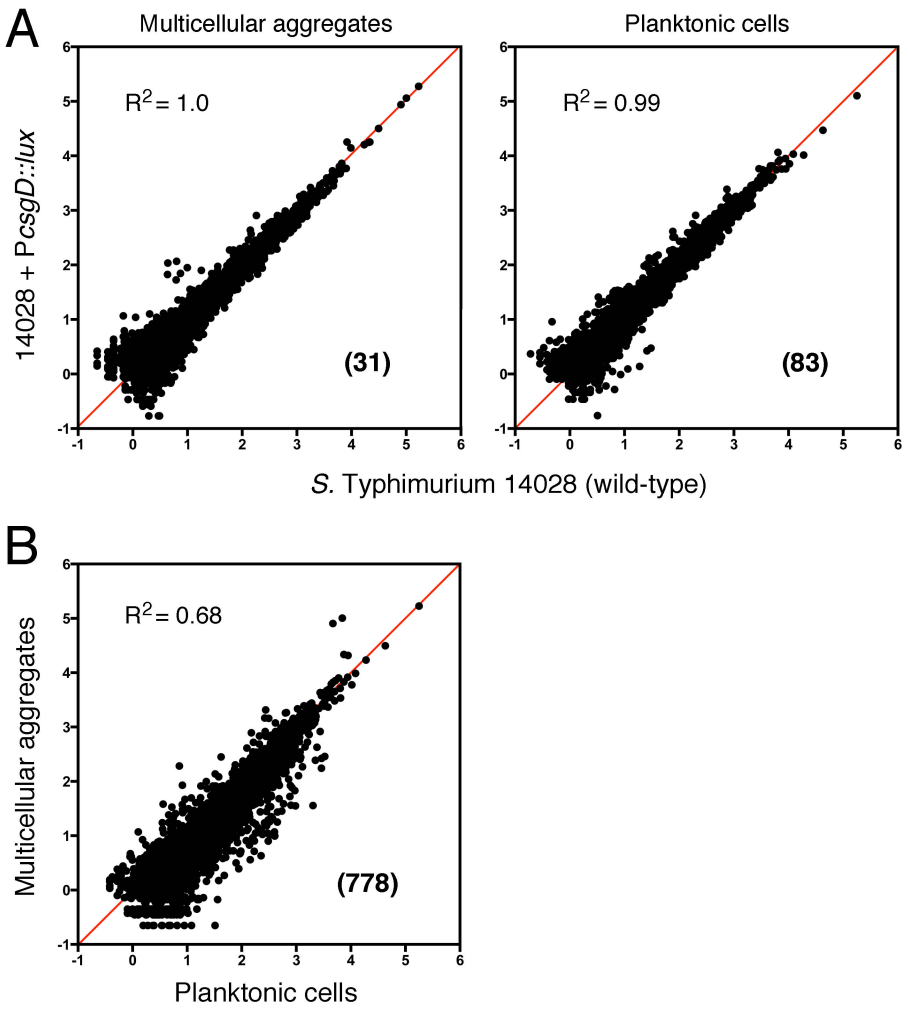

Figure 3.12. Scatterplot to compare RNA-seq transcriptome data from *S.* Typhimurium 14028 with and without a *csgD-luxCDABE* operon fusion plasmid.

RNA samples were purified from multicellular aggregates and planktonic cells isolated after 18 h growth in an *in vitro* flask model of biofilm development. Libraries were prepared and sequencing was performed on three biological replicate cultures. (A) For each gene, the average $\log_{10}$ read counts per million (CPM) values from *S.* Typhimurium 14028 cultures were plotted against the values obtained from *S.* Typhimurium 14028 + P*csgD::luxCDABE* cultures. The red line is the 45 degree line. The left panel represents multicellular aggregate samples and the right panel represents planktonic cell samples; the datasets are deposited in the SRA database (SRX976474 vs. SRX976343 for aggregates; SRX976470 vs. SRX976336 for planktonic cells). (B) This graph was included as a control to demonstrate a large difference between two sets of samples. For

each gene, the average $\log_{10}$ read CPM values obtained from multicellular aggregate samples vs. planktonic cell samples, prepared from *S.* Typhimurium 14028 cultures, are plotted. See reference 46 (MacKenzie *et al*., 2015) for a full description of this RNA-seq experiment, including data analysis. (A and B) R^2 values correspond to the two-tailed Pearson correlation coefficients. The bolded numbers in parentheses represent the total number of differentially expressed genes between the two datasets being compared (p <0.05).

by introducing a new antibiotic resistance marker in pHSG415, or perhaps through use of an alternative helper plasmid (17). Our modified Tn7 system was recently used to generate *luxCDABE* expressing strains of *S. enterica* serovar Enteritidis (S. Lam and W. Koester, unpublished) and *E. coli* K-12 (S. Bernier and M. Surette, unpublished), within one month of receipt of the plasmids. In theory, the use of the pHSG415-*tnsABCD* helper plasmid should allow for efficient Tn7 transposition in any bacterial species that can support replication of pSC101-based plasmids.

Reporter strains of *E. coli* and *S.* Typhimurium expressing bacterial luciferase or red or green fluorescent reporters were tested in combination with three synthetic, σ^{70}-dependent promoters. Due to the ubiquity of σ^{70} during growth (50), expression was constitutive and each promoter had similar temporal patterns that differed only in strength. The *S.* Typhimurium luciferase reporter strains provided stable expression throughout the infection of mice and three different levels of expression were observed. However, we were unable to detect *Salmonella* during the middle stages of infection (i.e., Day 2-4). It is possible that a stronger σ^{70}-dependent promoter, perhaps with both -10 and -35 consensus sequences (51), will allow visualization at lower cell numbers. The three levels of expression provide flexibility to ensure that the proper balance between reporter signal and minimal impact on biological processes can be achieved. It should be noted that there is a published method using Tn7 to generate bioluminescent *Salmonella* strains, and these authors demonstrated that luciferase output was proportional to cell number (21); however, their system does not provide for same flexibility that the pCS26-based system does. We have used pCS26 in the past to track the expression of various genes *in vivo* (32), but plasmid instability was a problem. The stochasticity of plasmid loss makes it difficult to form strong conclusions about the expression of different genes or to correlate with a specific number of bacterial cells. Now, with a system to easily integrate pCS26 constructs into the chromosome, expanded *in vivo* gene tracking experiments can be performed.

The *E. coli* and *S.* Typhimurium strains expressing fluorescent reporters were designed for future single cell applications, for example to analyze the dynamics within bacterial populations (52) or to study host-pathogen interactions (53). We demonstrated that these reporter strains could be detected with a 96-well plate reader, on solid media

using a whole animal imager and by confocal microscopy. Although there was differentiation between the σ^{70}-dependent promoters, the dynamic range of the fluorescent reporters was far less than luciferase, which is consistent with previous studies comparing these reporters (54). GFP and mCherry can be used for whole animal imaging experiments, but the autofluorescence from tissue and absorbance from specific cell molecules (i.e., hemoglobin) makes detection difficult (55). As a result, there has been a continual push to generate fluorescent reporters that have a lower background for whole animal imaging, such as those that emit signals in the near infrared range (i.e., >700nm (56)). If new reporters do become available (9), they can easily be incorporated into the pCS26 system and used to generate new chromosomal reporter strains.

The use of both kanamycin and chloramphenicol resistance markers facilitates competition experiments between *Salmonella* or *E. coli* strains or between different bacterial species. We had initial concerns with use of the kanamycin marker for infection experiments, due to the potential overgrowth of normal flora isolates upon recovery, but we did not encounter any problems. If a researcher does encounter low-level resistance, our tests showed that the Kn^R and Cm^R strains could be recovered equally well at higher antibiotic concentrations. We also generated a tetracycline-resistant pCS26 derivative, but luciferase expression was much higher and appeared to be unregulated (D.J. Shivak and A.P. White, unpublished data). It is assumed that insertion of transcriptional terminators upstream and downstream of the *tetC* gene would correct this problem, but we did not investigate this further.

Since the original publication of the Tn7 transposition method (12), it has proven to be an incredibly useful system. Modifications have been made in an attempt to improve the flexibility of the system (as described in the Introduction) and to incorporate new aspects of gene expression analysis (57). The work we have described will add to the microbiologist toolbox for researchers who study bacterial species within the *Enterobacteriaceae*, by providing great flexibility and the ability to combine with high-throughput expression experiments. The *tnsABCD* helper and miniTn7T delivery plasmids that we have generated will allow the development of chromosomal reporter strains in as little time as 2-4 weeks. These plasmids can be easily modified and are freely available upon request.

3.7 Acknowledgements

This research was supported by a grant from the Natural Sciences and Engineering Research Council (NSERC) (#386063-2010) to APW and through the Jarislowsky Chair in Biotechnology. KDM was supported by a Canada Graduate Scholarship from NSERC, DJS and NW by Research Awards from the University of Saskatchewan.

The authors are grateful to Don Wilson, Stew Walker and the VIDO animal care staff for professional help with the animal experiments; to Melissa Palmer, George Mutwiri and Elizabeth Brockman for laboratory assistance; and to Wei Xiao, Wolfgang Koester, and Shirley Lam and for helpful discussions.

3.8 References

1. **Baba T, Ara T, Hasegawa M, Takai Y, Okumura Y, Baba M, Datsenko KA, Tomita M, Wanner BL, Mori H.** 2006. Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. Mol Syst Biol **2:**2006 0008.

2. **Porwollik S, Santiviago CA, Cheng P, Long F, Desai P, Fredlund J, Srikumar S, Silva CA, Chu W, Chen X, Canals R, Reynolds MM, Bogomolnaya L, Shields C, Cui P, Guo J, Zheng Y, Endicott-Yazdani T, Yang HJ, Maple A, Ragoza Y, Blondel CJ, Valenzuela C, Andrews-Polymenis H, McClelland M.** 2014. Defined single-gene and multi-gene deletion mutant collections in *Salmonella enterica* sv Typhimurium. PLoS One **9:**e99820.

3. **Santiviago CA, Reynolds MM, Porwollik S, Choi SH, Long F, Andrews-Polymenis HL, McClelland M.** 2009. Analysis of pools of targeted *Salmonella* deletion mutants identifies novel genes affecting fitness during competitive infection in mice. PLoS Pathog **5:**e1000477.

4. **Kalir S, McClure J, Pabbaraju K, Southward C, Ronen M, Leibler S, Surette MG, Alon U.** 2001. Ordering genes in a flagella pathway by analysis of expression kinetics from living bacteria. Science **292:**2080-2083.

5. **Zaslaver A, Mayo AE, Rosenberg R, Bashkin P, Sberro H, Tsalyuk M, Surette MG, Alon U.** 2004. Just-in-time transcription program in metabolic pathways. Nat Genet **36:**486-491.

6. **Mills E, Petersen E, Kulasekara BR, Miller SI.** 2015. A direct screen for c-di-GMP modulators reveals a *Salmonella* Typhimurium periplasmic L-arginine-sensing pathway. Sci Signal **8:**ra57.

7. **Chaudhuri RR, Morgan E, Peters SE, Pleasance SJ, Hudson DL, Davies HM, Wang J, van Diemen PM, Buckley AM, Bowen AJ, Pullinger GD, Turner DJ, Langridge GC, Turner AK, Parkhill J, Charles IG, Maskell DJ, Stevens MP.** 2013. Comprehensive assignment of roles for *Salmonella typhimurium* genes in intestinal colonization of food-producing animals. PLoS Genet **9:**e1003456.

8. **Kroger C, Colgan A, Srikumar S, Handler K, Sivasankaran SK, Hammarlof DL, Canals R, Grissom JE, Conway T, Hokamp K, Hinton JC.** 2013. An infection-relevant transcriptomic compendium for *Salmonella enterica* Serovar Typhimurium. Cell Host Microbe **14:**683-695.

9. **Campbell-Valois FX, Sansonetti PJ.** 2014. Tracking bacterial pathogens with genetically-encoded reporters. FEBS Lett **588:**2428-2436.

10. **van Opijnen T, Camilli A.** 2013. Transposon insertion sequencing: a new tool for systems-level analysis of microorganisms. Nat Rev Microbiol **11:**435-442.

11. **Angelichio MJ, Camilli A.** 2002. In vivo expression technology. Infect Immun **70:**6518-6523.

12. **Choi KH, Gaynor JB, White KG, Lopez C, Bosio CM, Karkhoff-Schweizer RR, Schweizer HP.** 2005. A Tn7-based broad-range bacterial cloning and expression system. Nat Methods **2:**443-448.

13. **Choi KH, Schweizer HP.** 2006. mini-Tn7 insertion in bacteria with single *att*Tn7 sites: example *Pseudomonas aeruginosa*. Nat Protoc **1:**153-161.

14. **Peters JE, Craig NL.** 2001. Tn7: smarter than we thought. Nat Rev Mol Cell Biol **2:**806-814.

15. **Waddell CS, Craig NL.** 1989. Tn7 transposition: recognition of the *att*Tn7 target sequence. Proc Natl Acad Sci U S A **86:**3958-3962.

16. **Mitra R, McKenzie GJ, Yi L, Lee CA, Craig NL.** 2010. Characterization of the TnsD-*att*Tn7 complex that promotes site-specific insertion of Tn7. Mob DNA **1:**18.

17. **Crepin S, Harel J, Dozois CM.** 2012. Chromosomal complementation using Tn7 transposon vectors in *Enterobacteriaceae*. Appl Environ Microbiol **78:**6001-6008.

18. **McKenzie GJ, Craig NL.** 2006. Fast, easy and efficient: site-specific insertion of transgenes into enterobacterial chromosomes using Tn7 without need for selection of the insertion event. BMC Microbiol **6:**39.

19. **Damron FH, McKenney ES, Schweizer HP, Goldberg JB.** 2013. Construction of a broad-host-range Tn7-based vector for single-copy P(BAD)-controlled gene expression in gram-negative bacteria. Appl Environ Microbiol **79:**718-721.

20. **Damron FH, McKenney ES, Barbier M, Liechti GW, Schweizer HP, Goldberg JB.** 2013. Construction of mobilizable mini-Tn7 vectors for bioluminescent detection of gram-negative bacteria and single-copy promoter lux reporter analysis. Appl Environ Microbiol **79:**4149-4153.

21. **Howe K, Karsi A, Germon P, Wills RW, Lawrence ML, Bailey RH.** 2010. Development of stable reporter system cloning luxCDABE genes into chromosome of *Salmonella enterica* serotypes using Tn7 transposon. BMC Microbiol **10:**197.

22. **Bjarnason J, Southward CM, Surette MG.** 2003. Genomic profiling of iron-responsive genes in *Salmonella enterica* serovar Typhimurium by high-throughput screening of a random promoter library. J Bacteriol **185:**4973-4982.

23. **White AP, Weljie AM, Apel D, Zhang P, Shaykhutdinov R, Vogel HJ, Surette MG.** 2010. A global metabolic shift is linked to *Salmonella* multicellular development. PLoS One **5:**e11814.

24. **Zaslaver A, Bren A, Ronen M, Itzkovitz S, Kikoin I, Shavit S, Liebermeister W, Surette MG, Alon U.** 2006. A comprehensive library of fluorescent transcriptional reporters for *Escherichia coli*. Nat Methods **3:**623-628.

25. **Johnson TJ, Kariyawasam S, Wannemuehler Y, Mangiamele P, Johnson SJ, Doetkott C, Skyberg JA, Lynne AM, Johnson JR, Nolan LK.** 2007. The genome sequence of avian pathogenic *Escherichia coli* strain O1:K1:H7 shares strong similarities with human extraintestinal pathogenic *E. coli* genomes. J Bacteriol **189:**3228-3236.

26. **White AP, Gibson DL, Kim W, Kay WW, Surette MG.** 2006. Thin aggregative fimbriae and cellulose enhance long-term survival and persistence of *Salmonella*. J Bacteriol **188:**3219-3227.

27. **Herrero M, de Lorenzo V, Timmis KN.** 1990. Transposon vectors containing non-antibiotic resistance selection markers for cloning and stable chromosomal insertion of foreign genes in gram-negative bacteria. J Bacteriol **172:**6557-6567.

28. **White AP, Surette MG.** 2006. Comparative Genetics of the rdar Morphotype in *Salmonella*. Journal of Bacteriology **188:**8395-8406.

29. **Westerlund-Karlsson A, Saviranta P, Karp M.** 2002. Generation of thermostable monomeric luciferases from Photorhabdus luminescens. Biochem Biophys Res Commun **296:**1072-1076.

30. **Mitchell RJ, Ahn J, Gu MB.** 2005. Comparison of *Photorhabdus luminescens* and *Vibrio fischeri* lux fusions to study gene expression patterns. J Microbiol Biotechnol **15:**48-54.

31. **Kim W, Surette MG.** 2004. Metabolic differentiation in actively swarming *Salmonella*. Mol Microbiol **54:**702-714.

32. **White AP, Gibson DL, Grassl GA, Kay WW, Finlay BB, Vallance BA, Surette MG.** 2008. Aggregation via the red, dry, and rough morphotype is not a virulence adaptation in *Salmonella enterica* serovar Typhimurium. Infection and Immunity **76:**1048-1058.

33. **Hashimoto-Gotoh T, Franklin FC, Nordheim A, Timmis KN.** 1981. Specific-purpose plasmid cloning vectors. I. Low copy number, temperature-sensitive, mobilization-defective pSC101-derived containment vectors. Gene **16:**227-235.

34. **Choi KH, Mima T, Casart Y, Rholl D, Kumar A, Beacham IR, Schweizer HP.** 2008. Genetic tools for select-agent-compliant manipulation of *Burkholderia pseudomallei*. Appl Environ Microbiol **74:**1064-1075.

35. **Maloy SR, Stewart VJ, Taylor RK.** 1996. Genetic analysis of pathogenic bacteria: a laboratory manual. Cold Spring Harbor Laboratory Press, Plainview, NY.

36. **Livak KJ, Schmittgen TD.** 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods **25:**402-408.

37. **Barthel M, Hapfelmeier S, Quintanilla-Martinez L, Kremer M, Rohde M, Hogardt M, Pfeffer K, Russmann H, Hardt WD.** 2003. Pretreatment of mice with streptomycin provides a *Salmonella enterica* serovar Typhimurium colitis model that allows analysis of both pathogen and host. Infect Immun **71:**2839-2858.

38. **Chang AC, Cohen SN.** 1978. Construction and characterization of amplifiable multicopy DNA cloning vehicles derived from the P15A cryptic miniplasmid. J Bacteriol **134:**1141-1156.

39. **Datsenko KA, Wanner BL.** 2000. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. Proceedings of the National Academy of Sciences of the United States of America **97:**6640-6645.

40. **Kim W, Surette MG.** 2006. Coordinated regulation of two independent cell-cell signaling systems and swarmer differentiation in *Salmonella enterica* serovar Typhimurium. Journal of bacteriology **188:**431-440.

41. **Keestra-Gounder AM, Tsolis RM, Baumler AJ.** 2015. Now you see me, now you don't: the interaction of *Salmonella* with innate immune receptors. Nat Rev Microbiol **13:**206-216.

42. **Steenackers H, Hermans K, Vanderleyden J, De Keersmaecker SCJ.** 2012. *Salmonella* biofilms: An overview on occurrence, structure, regulation and eradication. Food Research International **45:**502-531.

43. **Romling U, Sierralta WD, Eriksson K, Normark S.** 1998. Multicellular and aggregative behaviour of *Salmonella typhimurium* strains is controlled by mutations in the agfD promoter. Molecular Microbiology **28:**249-264.

44. **Zogaj X, Nimtz M, Rohde M, Bokranz W, Romling U.** 2001. The multicellular morphotypes of *Salmonella typhimurium* and *Escherichia coli* produce cellulose as the second component of the extracellular matrix. Mol Microbiol **39:**1452-1463.

45. **Grantcharova N, Peters V, Monteiro C, Zakikhany K, Römling U.** 2010. Bistable expression of CsgD in biofilm development of *Salmonella enterica* serovar typhimurium. Journal of Bacteriology **192:**456-466.

46. **MacKenzie KD, Wang Y, Shivak DJ, Wong CS, Hoffman LJ, Lam S, Kroger C, Cameron AD, Townsend HG, Koster W, White AP.** 2015. Bistable expression of CsgD in *Salmonella enterica* serovar Typhimurium connects virulence to persistence. Infect Immun **83:**2312-2326.

47. **Arena ET, Auweter SD, Antunes LC, Vogl AW, Han J, Guttman JA, Croxen MA, Menendez A, Covey SD, Borchers CH, Finlay BB.** 2011. The deubiquitinase activity of the *Salmonella* pathogenicity island 2 effector, SseL, prevents accumulation of cellular lipid droplets. Infect Immun **79:**4392-4400.

48. **Sham HP, Shames SR, Croxen MA, Ma C, Chan JM, Khan MA, Wickham ME, Deng W, Finlay BB, Vallance BA.** 2011. Attaching and effacing bacterial effector NleC suppresses epithelial inflammatory responses by inhibiting NF-kappaB and p38 mitogen-activated protein kinase activation. Infect Immun **79:**3552-3562.

49. **Caulcott CA, Dunn A, Robertson HA, Cooper NS, Brown ME, Rhodes PM.** 1987. Investigation of the effect of growth environment on the stability of low-copy-number plasmids in *Escherichia coli*. J Gen Microbiol **133:**1881-1889.

50. **Gruber TM, Gross CA.** 2003. Multiple sigma subunits and the partitioning of bacterial transcription space. Annual Review of Microbiology **57:**441-466.

51. **Gaal T, Ross W, Estrem ST, Nguyen LH, Burgess RR, Gourse RL.** 2001. Promoter recognition and discrimination by EsigmaS RNA polymerase. Mol Microbiol **42:**939-954.

52. **Locke JC, Elowitz MB.** 2009. Using movies to analyse gene circuit dynamics in single cells. Nat Rev Microbiol **7:**383-392.

53. **Avraham R, Haseley N, Brown D, Penaranda C, Jijon HB, Trombetta JJ, Satija R, Shalek AK, Xavier RJ, Regev A, Hung DT.** 2015. Pathogen Cell-to-Cell Variability Drives Heterogeneity in Host Immune Responses. Cell **162:**1309-1321.

54. **de Jong H, Ranquet C, Ropers D, Pinel C, Geiselmann J.** 2010. Experimental and computational validation of models of fluorescent and luminescent reporter genes in bacteria. BMC Syst Biol **4:**55.

55. **Troy T, Jekic-McMullen D, Sambucetti L, Rice B.** 2004. Quantitative comparison of the sensitivity of detection of fluorescent and bioluminescent reporters in animal models. Mol Imaging **3:**9-23.

56. **Filonov GS, Piatkevich KD, Ting LM, Zhang J, Kim K, Verkhusha VV.** 2011. Bright and stable near-infrared fluorescent protein for in vivo imaging. Nat Biotechnol **29:**757-761.

57. **Bruckbauer ST, Kvitko BH, Karkhoff-Schweizer RR, Schweizer HP.** 2015. Tn5/7-lux: a versatile tool for the identification and capture of promoters in gram-negative bacteria. BMC Microbiol **15:**17.

4.0 BISTABLE EXPRESSION OF CsgD IN *SALMONELLA* CONNECTS VIRULENCE TO PERSISTENCE

Keith D. MacKenzie[1,2], Yejun Wang[1], Dylan J. Shivak[1,2], Cynthia S. Wong[1], Leia J.L. Hoffman[1], Shirley Lam[1], Carsten Kröger[3], Andrew D. S. Cameron[4], Hugh G.G. Townsend[1,5], Wolfgang Köster[1,6], and Aaron P. White[1,2]*

1 Vaccine and Infectious Disease Organization-International Vaccine Centre, Saskatoon, Saskatchewan, Canada[a],

2 Department of Microbiology and Immunology, University of Saskatchewan, Saskatoon, Saskatchewan, Canada[b],

3 Institute of Integrative Biology, University of Liverpool, Liverpool, United Kingdom[c],

4 Department of Biology, University of Regina, Regina, Saskatchewan, Canada[d],

5 Department of Large Animal Clinical Sciences, University of Saskatchewan, Saskatoon, Saskatchewan, Canada[e],

6 Department of Veterinary Microbiology, University of Saskatchewan, Saskatoon, Saskatchewan, Canada[f]

* Corresponding author, e-mail: aaron.white@usask.ca

Full citation: MacKenzie KD, Wang Y, Shivak DJ, Wong CS, Hoffman LJ, Lam S, Kroger C, Cameron AD, Townsend HG, Koster W, White AP. 2015. Bistable expression of CsgD in *Salmonella enterica* serovar Typhimurium connects virulence to persistence. Infect Immun **83:**2312-2326.

Author contributions:

Keith MacKenzie was responsible for experimental design, data generation, and for authoring the manuscript. He assisted with statistical analyses, with exception of the statistical analysis for the RNA-seq experiments.

Yejun Wang provided the bioinformatics analysis for the RNA-seq data and authored the manuscript.

Dylan Shivak created the kanamycin- and chloramphenicol-marked *Salmonella* strains, provided assistance with the competitive index assays, and edited the manuscript.

Cynthia Wong performed the western blotting assays.

Leia Hoffman assisted with the time point determination assays and with RNA isolation and purification.

Shirley Lam assisted with maintaining the Caco-2 cell line and with the Caco-2 invasion assays.

Carsten Kröger and **Andrew Cameron** are credited for preparing and sequencing the RNA-seq libraries for the Δ*csgD* mutant strain.

Hugh Townsend assisted with both the design and statistical analysis of the competitive infection assays.

Wolfgang Köster provided guidance with experimental design and edited the manuscript.

Aaron White assisted with experimental design and statistical analyses. He performed the swim assays, authored the manuscript, and held the research grant that supported this work.

4.1 Interface

This chapter presents our molecular and phenotypic analysis of the multicellular aggregates and planktonic cells that arise due to bistable expression of *csgD*. The results presented in this chapter provided the first published characterization of these subpopulations. Our focus was to understand the consequences of the CsgD bistability phenomenon that was previously described in the literature. As such, we designed our experiments to directly compare the subpopulations of multicellular aggregates and planktonic cells. We determined that the multicellular aggregates have a gene expression profile consistent with cell survival, whereas planktonic cells express multiple genes associated with virulence. The kanamycin- and chloramphenicol-resistant *Salmonella* strains generated using the Tn7 transposition system in the previous chapter allowed us to directly compare the relative fitness of the subpopulations in assays of virulence and persistence and provided powerful functional validation of the gene expression profiles that we had identified for multicellular aggregates and planktonic cells.

4.2 Abstract

Pathogenic bacteria often need to survive in the host and the environment, and it is not well understood how cells transition between these equally challenging situations. For the human and animal pathogen *Salmonella enterica* serovar Typhimurium, biofilm formation is correlated with persistence outside a host but the connection to virulence is unknown. In this study, we analyzed multicellular aggregates and planktonic cell subpopulations that co-exist when *S.* Typhimurium is grown under biofilm-inducing conditions. These cell types arise due to bistable expression of CsgD, the central biofilm regulator. Despite being exposed to the same stresses, the two cell subpopulations had 1856 genes that were differentially expressed, as determined by RNA-seq. Aggregated cells displayed the characteristic gene expression of biofilms, whereas planktonic cells had enhanced expression of numerous virulence genes. Increased type three secretion synthesis in planktonic cells correlated with enhanced invasion of a human intestinal cell line and significantly increased virulence in mice as compared to the aggregates. However, when these same groups of cells were exposed to desiccation, the aggregates

survived better and the competitive advantage of planktonic cells was lost. We hypothesize that CsgD-based differentiation is a form of bet hedging, with single cells primed for host cell invasion and aggregated cells adapted for persistence in the environment. This would allow *S.* Typhimurium to spread the risks of transmission and ensure a smooth transition between the host and the environment.

4.3 Introduction

Salmonella persistence, or the ability of cells to survive in stressful conditions, has traditionally been studied separately from virulence. The continued success of these pathogens, however, dictates that both traits must be inherently connected. Most of the >2000 serovars of *Salmonella* collectively account for approximately 94 million annual cases of human gastroenteritis (1), a disease characterized by intestinal inflammation and diarrhea (2). The gastroenteritis-causing serovars are collectively referred to as non-typhoidal *Salmonella* (NTS), and serovar Typhimurium is one of the most commonly isolated in the world. Most NTS isolates are considered to be host-generalists with a cyclical lifestyle involving exposure to the environment in between host colonization events (3). This is in contrast to typhoidal *Salmonella* serovars (i.e., Typhi and Paratyphi A), which are human-restricted (4, 5). Despite their differences, both NTS and typhoidal *Salmonella* rely on type three secretion systems (T3SS) to stimulate invasion, uptake and survival within human epithelial cells and macrophages (6). While we have an improving understanding of pathogenesis, less is known about the strategies that pathogenic *Salmonella* strains employ to survive outside the host (7, 8). It is generally assumed that people become infected with *Salmonella* through ingestion of contaminated food and water, but there are also numerous examples of sporadic human cases where the causative source is unknown (9).

One strategy for bacterial survival is through the formation of biofilms, where cells aggregate and become embedded in a self-produced extracellular matrix, usually in contact with a surface. This has been described in *Salmonella* as the rdar (red, dry and rough) morphotype (10, 11). The extracellular matrix surrounding rdar-positive cells is comprised of protein polymers (curli fimbriae) and extracellular polysaccharides (cellulose and an O-antigen capsule) that physically link cells together (10-13). The genes

coding for these polymers are conserved throughout *Salmonella* (14, 15) and *Escherichia coli* (16, 17). Comparisons between rdar-positive and rdar-negative *S.* Typhimurium and *E. coli* strains have shown that embedded cells have enhanced persistence upon exposure to conditions of starvation, desiccation and treatment with disinfectants (15, 18-21), as well as an increased ability to adhere to both biotic and abiotic surfaces (22, 23). Rdar-positive cells can also cause infections after long time periods (24), but do not appear to be specifically adapted for virulence (25). Rdar-negative *Salmonella* and *E. coli* isolates tend to cause more invasive forms of disease (26, 27).

Most NTS isolates are rdar-positive and most typhoidal *Salmonella* isolates are rdar-negative (21, 25, 26). One possible explanation for this correlation is that the rdar morphotype is an adaptation necessary for life outside the host. However, there have been recent links between the rdar morphotype, immune stimulation and invasion (28, 29), making it difficult to fully understand this physiology. Although they are informative, a potential problem with strain-to-strain comparisons is that they could miss fluctuations that exist within individual populations of cells. It was recently demonstrated that when *S.* Typhimurium is grown under biofilm-inducing conditions in liquid media, the population of cells in the fluid phase differentiates into large, multicellular aggregates and single (planktonic) cells (25, 30). This bifurcation is caused by the bistable synthesis of CsgD, the central transcriptional regulator of the rdar morphotype (31). Aggregated cells are in a CsgD-ON state, whereas the non-aggregative planktonic cells are primarily in a CsgD-OFF state. The multicellular aggregates can be thought of as non-canonical biofilms, as the cells highly express extracellular matrix components (25, 30), but are not in direct contact with a surface. While it is assumed that the aggregates retain the resistance properties of classic adherent biofilms, the planktonic cells that result from CsgD bistability have never been analyzed in detail and it is unclear what role they may have in *S.* Typhimurium physiology.

In this study, we have isolated and examined these two groups of cells and demonstrate for the first time that they could represent a critical link between the virulence and persistence of *Salmonella*. Bistability and the formation of specialized cell types are predicted to contribute to *S.* Typhimurium transmission success, with single

cells adapted for host cell invasion and multicellular aggregates adapted for persistence in the environment.

4.4 Materials and Methods

4.4.1 Bacterial strains, media, and growth conditions.

Salmonella enterica serovar Typhimurium strain ATCC 14028 was the wildtype strain in this study. In general, strains were inoculated from frozen stocks onto LB agar and grown for 20 h at 37˚C. One isolated colony was used to inoculate 5 mL LB broth (containing 1% NaCl) and the culture was incubated for 12 h at 37˚C with agitation at 200 rpm. To analyze growth under biofilm-inducing conditions, approximately 4 x 10^9 CFU were inoculated into 400 mL of 1% tryptone pH 7.4 and the culture was distributed in 100 mL volumes into four 250 mL flasks that were incubated at 28˚C for 8h, 13h, 18 h, or 32 h with agitation at 200 rpm. The *csgD* promoter (P*csgD*) *luxCDABE* fusion plasmid used for bioluminescence assays has been previously described (21). To measure luminescence, 200 µl aliquots of culture were removed at hourly intervals in triplicate using a wide-bore pipette, transferred into individual wells of a 96-well clear-bottom black plate (Product #9520, Corning Life Sciences), and assayed on a Victor X^3 multilabel plate reader (Perkin-Elmer Life Sciences).

4.4.2 Generation of *S.* Typhimurium 14028 Δ*csgD* and ΔSPI-1 mutant strains.

Δ*csgD* and Δ*SPI-1* strains were generated using the lambda red recombinase knockout procedure (32). Primers containing a 50 nucleotide sequence on either side of the entire SPI-1 region (SPI-1_F and SPI-1_R) (33) or *csgD* (csgDko sense; GCAGCTGTCAGATGTGCGATTAAAAAAAGTGGAGTTTCATCATGTTTAATGT GTAGGCTGTAGCTGCTTC and csgDko antisense; CTCTGCTGCTACAATCCAGG TCAGATAGCGTTTCATGGCCTTACCGCCTGCCTCCTTAGTTCCTATTCCG) were used to amplify the *cat* gene from pKD3 using Phusion High-Fidelity DNA polymerase (New England Biolabs). PCR products were purified (AxyPrep Mag PCR clean-up kit; Axygen) and electroporated into *S.* Typhimurium 14028 cells containing pKD46. Mutant strains were first selected by growth at 37°C on LB agar supplemented with 7 µg mL^{-1}

Cm, before re-streaking onto LB agar containing 30 μg mL^{-1} Cm. Mutations were moved into clean wild-type backgrounds by transduction with P22 phage (34). The *cat* gene was resolved from the chromosome using pCP20 (32). DNA sequencing of PCR products amplified from the chromosome of Δ*csgD* or ΔSPI-1 strains confirmed the loss of the majority of *csgD* or SPI-1, respectively.

4.4.3 Generation of Kn^R or Cm^R-resistant *S.* Typhimurium strains.

sig70-16 is a reporter plasmid that contains a synthetic, σ^{70}-dependent promoter driving *luxCDABE* expression (35). A chloramphenicol-resistant (Cm^R) version of the sig70-16 plasmid was generated, and the sig70_16-Kn^R or -Cm^R fragments were used for Tn7 integration (36) into the *S.* Typhimurium ATCC 14028 chromosome, using a modified protocol (D. Shivak, K. MacKenzie, B. Jones and A. White, submitted for publication). To confirm that the fragments had been properly inserted into the *att*Tn7 site downstream of the *glmS* gene of *S.* Typhimurium, the chromosomal region was PCR amplified and the resulting products were sequenced. To ensure the absence of secondary mutations, the Kn^R and Cm^R sig70_16 fragments were moved into a clean wild-type or Δ*csgD* background by P22 transduction.

4.4.4 Isolation of multicellular aggregates and planktonic cell subpopulations.

S. Typhimurium biofilm cultures were transferred into 15 mL conical tubes and subjected to slow-speed centrifugation (210 x *g,* 2 min, 4°C). The supernatant fractions were transferred into new 15 mL tubes, and the cell aggregates were transferred into 1.5 mL tubes using a sterile loop. After brief centrifugation (10,000 x g, 30 s), culture media was removed and the wet weight of the aggregates was determined. Aggregates were resuspended in 1 mL of PBS (for direct counting) or 1 mL cell-free 1% tryptone culture supernatant (for subsequent applications) and homogenized for approximately 30 s (25 dounces) with a glass tissue grinder (Product #7727-2, Corning Life Sciences). For CFU determination, serial dilutions of planktonic cells or homogenized aggregates were plated in 4 μL drops onto LB agar, prior to incubation overnight at 37°C.

4.4.5 Visualizing the homogenization of multicellular aggregates

Multicellular aggregates were harvested from a *S.* Typhimurium 14028 reporter strain containing a *csgBAC* promoter - GFP fusion (25) grown for 13 h (TP2). 10 mg of aggregates were resuspended in 1 mL of cell-free culture supernatant, homogenized in a tissue grinder, and two successive 1 in 10 dilutions were used to prepare a solution of 1.2 x 10^7 CFU/mL. Ten 200 mL aliquots were individually concentrated onto the surfaces of 3 inch x 1 inch x 1.0 mm glass slides (Fisher Scientific #12-550-343) using a Cytospin-4 cytocentrifuge (Thermo Shandon, Asheville, NC). Samples were centrifuged at 500 rpm for 5 min with low acceleration. 5 mL of Mowiol mounting medium (Sigma-Aldrich # 81381) was added to the cells on each slide and each mixture was topped with a #1 ½ micro-cover glass (12 mm diameter) (Electron Microscopy Sciences #72230-01). Slides were air-dried overnight prior to visualization on a Leica TCS SP5 Confocal microscope, using the 63x oil immersion objective with the laser operating at 488 nm.

4.4.6 SDS-PAGE and western blotting.

Planktonic cells were sedimented (11,000 x g, 10 min, 4°C) to concentrate the samples, before suspending in 1 mL of SDS-PAGE sample buffer (without 2-mercaptoethanol and bromophenol blue). 100 mg samples of multicellular aggregates were resuspended in 1 mL of water and homogenized with a glass tissue grinder. Cell material was sedimented (10,000 x g, 1 min) before resuspending in 1 mL of SDS-PAGE sample buffer and boiling for 5 min. Protein concentrations were determined using the DC protein assay kit (BioRad Laboratories), and samples were normalized to 30 mg of total protein per lane and Bromophenol blue (0.0002% final conc.) and 2-mercaptoethanol (0.2% final conc.) were added. SDS-PAGE was performed with a 5% stacking gel and a 12% resolving gel or with 10%, 12% or 4-15% pre-cast Mini-PROTEAN TGX™ gels (Bio-Rad Laboratories). For immunoblots, proteins were transferred to nitrocellulose for 40 min at 25 V using a Trans-Blot SD semi-dry transfer cell (Bio-Rad Laboratories) in buffer recommended by the manufacturer. For detection of curli fimbriae, cell samples were boiled for 5 min in SDS-PAGE sample buffer with 2-mercaptoethanol, the cell debris was sedimented (14,000 x *g*, 2 min), washed twice in

500 mL of distilled water, before suspending in 250 mL of 90% formic acid, freezing at -80°C and lyophilization (10). Dried samples were resuspended in SDS-PAGE sample buffer and loaded directly, without boiling, into each gel lane prior to electrophoresis.

CsgD was detected using the monoclonal antibody described below. CsgA, the major subunit of curli fimbriae, was detected using rabbit polyclonal immune serum raised to whole purified curli (10). GroEL was used as a protein-loading control and was detected with rabbit polyclonal immune serum used at a 1 in 10,000 dilution (Sigma-Aldrich, G6532). SPI-1 TTSS proteins (InvG, SipD, SopE2, and SopB) were detected with rabbit polyclonal immune sera generated against each of the individual proteins (33). Secondary antibodies were goat anti-rabbit immunoglobulin G (IgG)- or human-adsorbed, goat-anti-mouse IgG-alkaline phosphatase conjugates (KPL; www.kpl.com). Proteins were detected by dye precipitation using the substrate 5-bromo-4- chloro-3-indolylphosphate and 4-nitroblue tetrazolium chloride (Sigma-Aldrich) as an enhancer.

4.4.7 Generation of CsgD-specific monoclonal antibodies.

csgD was PCR-amplified from *S.* Typhimurium 14028 genomic DNA using Phusion High-Fidelity DNA polymerase (New England Biolabs) and primers csgD_expFOR (GGCTCATATGTTTAATGAAGTCCATAG) and csgD_expREV2 (ATTTCTCGAGCCGCCTGAGATTATCGTTTG). The PCR product was digested with *Nde*I and *Xho*I and ligated into *Nde*I/*Xho*I-digested pET-21a (Novagen) using T4 DNA ligase (NEB), to generate a vector encoding CsgD with a 6xHis tag at the C-terminal end. *E. coli* BL21-DE3 + pET21a-*csgDHis* was grown at 37°C with agitation at 200 rpm in LB supplemented with 100 μg mL^{-1} ampicillin until an OD_{600} of 0.7 was reached; expression of CsgD-His was induced by adding IPTG to a final concentration of 1mM and growing for 2 h. Cells were harvested, lysed and CsgD-His was purified using a Ni-NTA agarose column (Qiagen), following procedures outlined by the manufacturer (Qiagen; The QIAexpressionist, Protocol 9). Approximately 5 mg of purified CsgD-His was used for Rapid Prime - Custom Monoclonal Antibody Development (ImmunoPrecise Antibodies Ltd., Victoria, BC, Canada). Concentrated tissue culture supernatants containing a CsgD-specific monoclonal antibody were used for western blotting experiments.

4.4.8 Measurement of c-di-GMP levels inside cells.

Planktonic cells and multicellular aggregates were harvested from 100 mL *S.* Typhimurium biofilm cultures grown for 8 h (TP1) or 13 h (TP2). Planktonic cell samples were prepared by centrifugation (10,000 x *g*, 10 min) and multicellular aggregate samples were homogenized as above in cell-free culture supernatant. Cyclic-di-GMP extractions were performed with 100 uL of 40% methanol-40% acetone in 0.1 N formic acid used for every 48 mg wet weight of cells (37). Cell samples were transferred into screw-capped tubes containing 0.2 g of 0.1 mm Zirconia beads and homogenization was performed in a high-speed mixer mill (Retsch; MM400) for 2 minutes at 30 Hz. After centrifugation (16,000 x *g*, 2 min), the supernatant was neutralized by the addition of 4 ml of 15% ammonium bicarbonate per 100 ml of sample. 50 mL aliquots of each sample were dried down in a speed-vacuum concentrator and residues were dissolved in 300 uL of 30% methanol followed by solid-phase extraction with Agilent OPT 30mg/mL reversed-phase cartridges (Agilent# 5982-3013) to remove lipids and other lipophilic compounds. The flow-through fractions were collected under a 3-5 psi vacuum and the cartridges were washed with 400 mL of 30% methanol. The flow-through fractions were collected and pooled for each sample and were dried down in a speed-vacuum concentrator. The residues were resuspended in 50 mL of 80% acetonitrile and 20 uL aliquots were injected into a Dionex UPLC 3000 system coupled to an AB Sciex 4000 QTRAP mass spectrometer, which was operated in the electrospray ionization and negative ion multiple-reaction monitoring (MRM) mode. The MRM transition for quantitation was 689.2/344.2. UPLC separation was carried out on a 10-cm long UPLC column, which was run in a hydrophilic interaction (HILIC) mode. The mobile phase was ammonium acetate – acetonitrile for binary gradient elution. The column temperature was 30^{o}C and the flow rate was 0.40 mL/min. A standard curve for cyclic-di-GMP (www.axxora.com; #BLG-C057) was prepared within a concentration range of 2 to 400 fmol/mL. The concentration of cyclic-di-GMP in each sample was calculated with external calibration. Since multicellular aggregates consist of both cells and extracellular matrix, a conversion factor was applied to all c-di-GMP concentrations. 1 mg wet weight of planktonic cells was equivalent to 6.94 x 10^{8} CFU, whereas 1 mg wet weight of

multicellular aggregates was equivalent to 1.73 x 10^8 CFU. This conversion was thought to represent a more accurate cell-to-cell comparison between samples.

4.4.9 RNA purification procedure.

Total RNA was isolated using a modified procedure based on the SV Total RNA Isolation Kit (Promega). At TP1, cells were sedimented by centrifugation (10,000 x *g*, 2 min, 4ºC) and resuspended in 1 mL of RNALater (Life Technologies). At later time points, cell subpopulations were first separated by slow-speed centrifugation, as described above. Planktonic cell fractions were sedimented as above and resuspended in 1 mL of RNALater (Life Technologies). Multicellular aggregate samples were resuspended in 1 mL of RNALater and homogenized for approximately 30 s using a tissue grinder. Cells in 1 mL of RNALater were sedimented by centrifugation (13,200 x *g*, 1 min) and resuspended in 700 µl of RNA Lysis Buffer (RLA). This cell suspension was transferred to a 1.5 mL screw-capped tube containing 0.2 g of 0.1 mm Zirconia beads, and homogenized in a high-speed mixer mill (Retsch; MM400) for 5 minutes at 30 Hz. Approximately 300 – 400 uL RLA solution was recovered from the beads, 600 – 800 mL of RNA Dilution Buffer was added, followed by 0.4 volumes of 95% ethanol. Samples were centrifuged (13,200 x *g*, 10 min) to remove extracellular debris that precipitated upon addition of ethanol, prior to loading onto the RNA purification column. This step was necessary to avoid plugging the column, especially with the multicellular aggregate samples. Column washing steps and on-column DNase I treatment were performed according to manufacturer's instructions. Bound RNA was eluted from each column with 2 x 50 µL of RNase-free water for a final volume of 100 µL. After elution, samples were DNase treated a second time in solution using a Turbo DNA-free kit (Life Technologies #AM1907). The MICROB*Express*™ Bacterial mRNA Enrichment Kit (Life Technologies #AM1905) was used for mRNA enrichment. To assess quantity, purity and integrity of RNA, samples were analyzed using a NanoDrop ND-1000 *s*pectrophotometer (Fisher Scientific) and Agilent 2100 Bioanalyzer with a Prokaryote Total RNA Nano chip.

4.4.10 Directional RNA-seq

For sequencing on the Illumina HiSeq 2000, first strand cDNA was synthesized from RNA using random hexamer priming and reverse transcriptase. Second strand cDNA synthesis was performed with DNA polymerase I in presence of dATP, dGTP, dCTP and dUTP. Double-stranded cDNA was fragmented by sonication and 70-200 bp fragments were purified with AMPure magnetic beads (Beckman-Coulter). After the standard addition of 18 mL of AMPure beads to every 10 mL of sample, 9.3 mL isopropanol was added. This step allows recovery of fragments below 100 bp, as described in the AxyPrep Mag PCR Clean-up kit protocol (Axygen Biosciences Inc.). cDNA fragments were ligated to Illumina adapters and primers containing bar-coded sequences; fragment sizes in the final library ranged from approximately 196 to 326 bp. The second cDNA strand containing UTP was selectively removed by uridine digestion (37°C, 15 min) in 5 mL of 1x TE buffer, pH 7.5 with 1 unit of Uracil-N-Glycosylase (UNG; Applied Biosystems), prior to library amplification and sequencing (38). In total, 21 libraries were sequenced (75 bp paired-end PET sequencing), prepared from three biological replicates of planktonic cells (samples at TP1, TP2, TP3 and TP4) and multicellular aggregates (samples at TP2, TP3 and TP4).

For sequencing on the Illumina MiSeq, cDNA libraries prepared from *S.* Typhimurium 14028 Δ*csgD* at TP2 were constructed using the NEBNext Ultra Directional RNA Library Prep Kit for Illumina (New England Biolabs) as per the manufacturer's protocol. This differs from the above protocol in that fragmentation occurs through autocatalysis of RNA before first strand cDNA synthesis. Final library fragment sizes were 200-300 bp. This approach also incorporates dUTP into the second cDNA strand synthesis followed by digestion of the second strand before PCR amplification. 75 bp paired-end PET sequencing was performed on three libraries using the MiSeq Reagent Kits v3 (Illumina).

4.4.11 Mapping and Comparison of Sequencing Reads

Sequencing reads were mapped onto the *S.* Typhimurium ATCC 14028 chromosome (NC_016856) and plasmid (NC_01685) sequence using Geneious v5.6.5

(www.biomatters.com). The mapping parameters were set as following: 1) gaps not allowed; 2) words repeated more than twice ignored; and 3) maximum mismatches per read set as 4%. Other parameters conform to the default settings of Geneious. The coordinates and strand of each putative protein-encoding gene were according to genome annotations (39), and in-house perl scripts were used for counting the number of mapping reads to each protein encoding gene and its antisense strand. The sRNA coverage was counted in a similar way. The full set of known *Salmonella* sRNAs (40), were retrieved and the orthologous counterparts in *S.* Typhimurium 14028 were found by direct sequence alignment.

Before comparing the gene expression levels between samples, the counts of mapped reads were normalized with edgeR, which, instead of using RPKM notation, assumes the total RNA amount could vary among different samples, and expression of more genes should be stable among libraries (41). A negative binomial distribution was fit for the normalized read counts for each pair of genes. Differentially expressed genes were identified as those with p value < 0.05, determined by a Benjamini & Hochberg correction for the multiple testing results (42, 43) with the False Discovery Rate (FDR) set as < 0.05. The normalization effect was evaluated before data analysis by observing the variance among technical and biological repeats after normalization.

4.4.12 Motility comparison between planktonic cells and multicellular aggregates.

Cell subpopulations were isolated from *S.* Typhimurium biofilm cultures at TP2 and either tested for motility directly or after homogenization for 30s in a tissue grinder. 1 mL aliquots (approximately 5 x 10^6 CFU) of each cell suspension were inoculated into the middle of swim agar (lysogeny broth (Luria) + 0.25% agar) and plates were allowed to dry for 20 min. Migration distances from the center of the plate were recorded after incubation at 37°C for 8 h. Five technical replicates of each sample were tested prepared from three biological replicate cultures. Statistical difference between samples was calculated using unpaired t tests (** p <0.01).

4.4.13 Invasion assay using polarized Caco-2 cells.

Caco-2 cells (obtained from Brett Finlay, University of British Columbia) were grown at 37°C in the presence of 5% CO_2 in DMEM complete media (HyClone Dulbecco modified Eagle medium (Fisher Scientific) supplemented with 10% fetal bovine serum (Seracare Life Sciences) and 1% nonessential amino acids (Life Technologies)). To obtain polarized monolayers, Caco-2 cells were seeded onto Transwell inserts (24 mm in diameter, 0.4-mm pore size; product #3470, Corning Life Sciences) for approximately 21 days. The cells were used for invasion assays once the transepithelial resistance (TER) was 700 - 900 Ω cm^{-2} (33). Planktonic cells and multicellular aggregates isolated from *S.* Typhimurium cultures were suspended in DMEM complete media and applied in 200 mL aliquots to the apical surface of the Caco-2 cells; invasion in three replicate wells was assessed for every sample. Wildtype and Δ*SPI-1* mutant strains grown to late-exponential phase, at an OD_{600} of 0.7 in LB at 37°C, were used as positive and negative invasion controls, respectively. After 1 h exposure to *S.* Typhimurium, Caco-2 cells were washed three times with 200 μL of PBS and incubated for 2 h with DMEM complete containing 1.2 mg mL^{-1} gentamicin to kill any remaining extracellular bacteria. Caco-2 cells were washed two times with 200 μl of PBS and lysed by exposure to 1% Triton. Serial dilutions of the *Salmonella*-containing lysate were plated in 4 mL drops on LB agar and incubated overnight at 37ºC. We aimed for a multiplicity of infection (MOI) of 10 based on an estimated 2.5 x 10^5 polarized Caco-2 cells per well. The true MOI was determined by serial dilution and plating on LB agar followed by incubation overnight at 37°C.

4.4.14 Gentamicin susceptibility testing.

Samples containing approximately 5 x 10^7 CFU mL^{-1} of planktonic cells or multicellular aggregates isolated from *S.* Typhimurium 14028 cultures were prepared in cell-free, spent culture media. 1 mL aliquots were centrifuged (8000 x *g*, 2 min), and cells were resuspended in 1 mL of DMEM complete containing gentamycin at concentrations ranging from 0.2 to 1.2 mg mL^{-1}. After 2 h exposure, cells were sedimented, washed and resuspended in PBS, and homogenized for 30s using a tissue grinder. Serial dilutions of each sample were plated on LB agar and grown overnight at 37°C to determine the

number of CFU. Four technical replicates of each sample were tested and the experiment was repeated three times. Statistical difference between samples was calculated using unpaired *t* tests (* $p < 0.05$, ns $p > 0.05$).

4.4.15 Competitive index (CI) infections of mice.

Female C57BL/6 mice (6 to 8 weeks old) purchased from Jackson Laboratory (Bar Harbor, ME) were assigned to cage groups using a randomization table prepared in Microsoft Excel and individual mice were marked with ear notches. Challenge formulations A and B were prepared in 100 mM HEPES pH8 at a concentration of ~10^7 CFU per 100 µL, consisting of an approximate 1:1 ratio of cells isolated from Cm^R or Kn^R *S.* Typhimurium cultures at TP2 (Fig. 4.11). Animal care technicians assigned challenge A and B to 2 cages each (i.e., 4 groups of 6 mice, $n = 24$) and performed all infections. Cage assignments were only revealed to members of the White lab after the final results were analyzed. Infected mice were weighed daily and monitored for clinical signs of infection; mice that had a >20% drop in weight were euthanized, typically 4 – 7 days post-infection. The spleen, liver, cecum and mesenteric lymph nodes (MLN) were homogenized in Eppendorf Safe-Lock Tubes containing 1 mL of PBS and a 5 mm stainless steel bead (Qiagen # 69989) using a high-speed mixer mill (Retsch; MM400). To determine the number of CFU in the initial challenges or organ homogenates, serial dilutions were plated on LB agar supplemented with 10 µg mL^{-1} Cm or 50 µg mL^{-1} Kn. The competitive index (CI) was calculated as $[\text{CFU planktonic/CFU biofilm}]_{output}/[\text{CFU planktonic/CFU biofilm}]_{input}$. For statistical analysis, the normality of each set of CI data was assessed using the Shapiro-Wilk normality test followed by the Wilcoxon signed rank test to determine if the median CI was significantly different from 1 (** $p<0.01$, *** $p < 0.001$, **** $p < 0.0001$, ns $p > 0.05$).

4.4.16 Input and output ratios for CI infections.

For the initial CI trial, the input ratio for challenge A was 0.76 (7.9×10^6 CFU planktonic cells, and 1.0×10^7 CFU aggregates) and for challenge B was 0.60 (5.9×10^6 CFU planktonic cells and 9.9×10^6 CFU aggregates). To analyze the effect of homogenization, challenge mixtures were homogenized in a tissue grinder for

approximately 30 s before oral gavage. The input ratio for challenge A was 1.19 (4.8×10^6 CFU planktonic cells and 4.0×10^6 CFU aggregates) and for challenge B was 0.85 (5.6×10^6 CFU planktonic cells and 6.6×10^6 CFU aggregates). Input ratios for the wildtype planktonic cells versus Δ*csgD* planktonic cells trial were 1.07 for challenge A (6.7×10^6 CFU wildtype, and 6.3×10^6 CFU Δ*csgD*) and 1.11 for challenge B (7.7×10^6 CFU wildtype, and 6.9×10^6 CFU Δ*csgD*). For the trial using planktonic cells and multicellular aggregates prepared from ΔSPI-1 cultures, the input ratio for challenge A was 0.27 (4.8×10^6 CFU planktonic cells, and 1.8×10^7 CFU aggregates) and for challenge B was 3.7 (4.4×10^6 CFU planktonic cells and 1.2×10^6 CFU aggregates). For the trial using dehydrated planktonic cells and multicellular aggregates (see below), the input ratio for challenge A was 0.11 (7.2×10^5 CFU planktonic cells, and 6.5×10^6 CFU aggregates) and for challenge B was 0.07 (3.0×10^5 CFU planktonic cells, and 4.2×10^6 CFU aggregates).

4.4.17 Desiccation of *S.* Typhimurium challenge formulations.

Planktonic cells and multicellular aggregates from the initial CI trial were suspended in spent culture supernatant and 200 mL aliquots were pipetted into individual wells of 24-well tissue culture plates. The plates were placed in a biological safety cabinet, the liquid was allowed to evaporate, and then plates were covered and left to sit on the lab bench. At time 0, 4 weeks and 5.5 weeks, 500 mL of PBS was added to individual wells and the cells were allowed to rehydrate for 30 min prior to homogenization using a tissue grinder and plating to determine the number of surviving CFU from each cell type. Statistical difference between samples was calculated using unpaired *t* tests (** $p < 0.01$, **** $p < 0.0001$). For the CI trial performed at 4 weeks, cells were resuspended in a total volume of 3 mL of 100 mM HEPES pH 8 per 24-well plate.

4.4.18 Ethics Statement

All animals were cared for and used in accordance with the Guidelines of the Canadian Council on Animal Care, the Regulations of the University of Saskatchewan Committee on Animal Care and Supply, and in accordance to the “3R principles”. All

mouse experiments were performed under Animal Use Protocol # 20110057 that was approved by the University of Saskatchewan's Animal Research Ethics Board.

4.5 Results

We grew *S.* Typhimurium 14028 in liquid culture under biofilm inducing conditions to examine differentiation within the population. After 13 h growth, multicellular aggregates were clearly visible within the culture (Fig. 4.1A), ranging in size between 100 – 500 mm (Fig. 4.1B). The aggregates had high levels of *csgBAC* (curli) transcription, as measured using a GFP reporter (Fig. 4.1C). To monitor the dynamics of the differentiation process, we measured the real-time expression of *csgD* in the total population while observing the appearance of multicellular aggregates (Fig. 4.2A). Four time points were chosen for further analysis: Time point 1 (TP1) was considered a pre-aggregation condition, when overall *csgD* expression was low (Fig. 4.2A) and the culture appeared homogeneous (data not shown); TP2 had peak *csgD* expression that coincided with the appearance of multicellular aggregates; TP3 had high *csgD* expression and the culture appeared fully differentiated; and TP4 was visually the same as TP3, but *csgD* expression had dropped to near baseline levels. Slow-speed centrifugation was used to sediment the multicellular aggregates while the planktonic cells remained suspended in the medium (Fig. 4.3). The formation of these two cell types was consistent, with a coefficient of variation of less than 10% between biological replicate cultures. The aggregates could be disrupted using a tissue grinder, which generated a primarily single cell population. Using this technique, we were able to establish a correlation between the wet weight of aggregated cells and the number of colony forming units (Fig. 4.3). In addition to the cells in the fluid phase, there were cells that formed an adherent biofilm at the air-liquid interface of the culture flask. These cells differ in terms of oxygen availability (30) and the amount of biofilm formed was highly variable (data not shown), therefore, they were excluded from further analysis.

Western blot analysis of the subpopulations isolated from the fluid phase confirmed the bistable production of CsgD; multicellular aggregates displayed high levels of CsgD at all time points, whereas the planktonic cells had either low (TP2), trace (TP1,

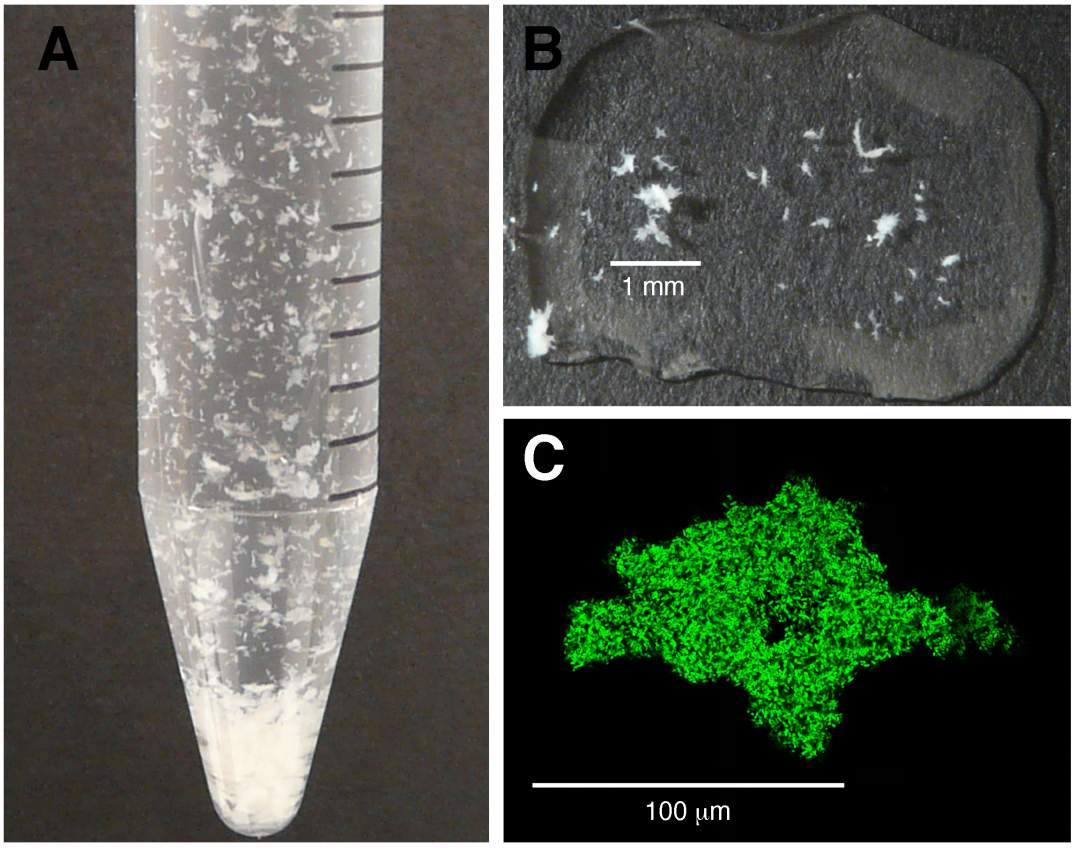

Figure 4.1. Differentiation of cells within a *S.* Typhimurium biofilm culture.

(A) An aliquot of *S.* Typhimurium 14028 culture containing multicellular aggregates and planktonic cells is shown after 13h growth at 28ºC, prior to separation of the two subpopulations. (B) Isolated aggregates are shown resuspended in PBS to give an example of their relative size. (C) Aggregates were isolated from a *S.* Typhimurium culture containing a *csgBAC* promoter-GFP fusion plasmid (25) and visualized on a Leica TCS SP5 Confocal microscope.

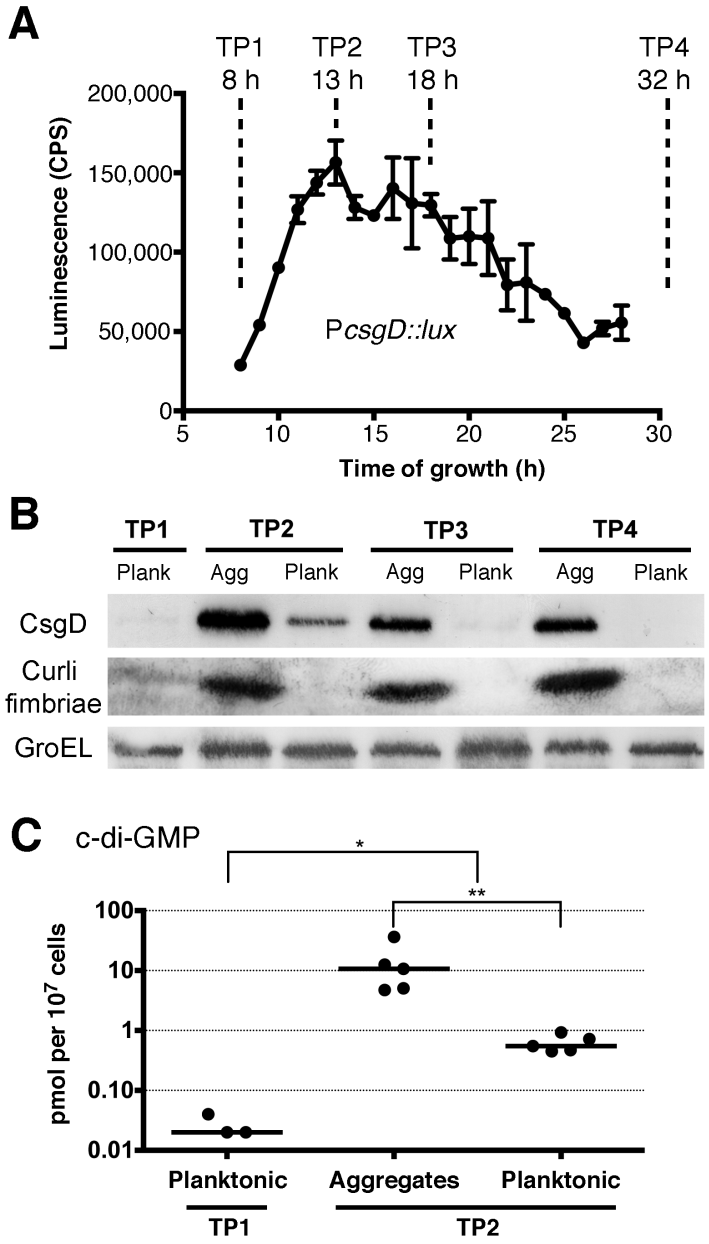

Figure 4.2. Analysis of multicellular aggregates and planktonic cells isolated from *S.* Typhimurium biofilm cultures.

(A) *csgD* expression was measured during growth using a promoter-luciferase (*luxCDABE*) fusion plasmid, designed to measure gene expression by light production (counts per second, CPS). The line represents the average CPS of two biological replicate cultures, measured at hourly intervals in triplicate; the error bars represent the range of expression. Time points (TP1-4) were chosen to reflect the dynamics of cellular differentiation. (B) Cell fractions were isolated at each time point and analyzed by

western blotting for synthesis of CsgD, the central regulator of *Salmonella* biofilm formation, and CsgA, the major subunit of curli fimbriae. GroEL was detected as a loading control to ensure that equal amounts of protein were loaded into each lane. (D) Cyclic-di-GMP concentration values were calculated as pmol per 10^7 CFU. Each point represents the measurement from one biological replicate culture; c-di-GMP could not be detected for two samples of planktonic cells at TP1. The horizontal bars represent the median values. Statistical significance is noted as: *, $p < 0.05$, **, $p < 0.01$.

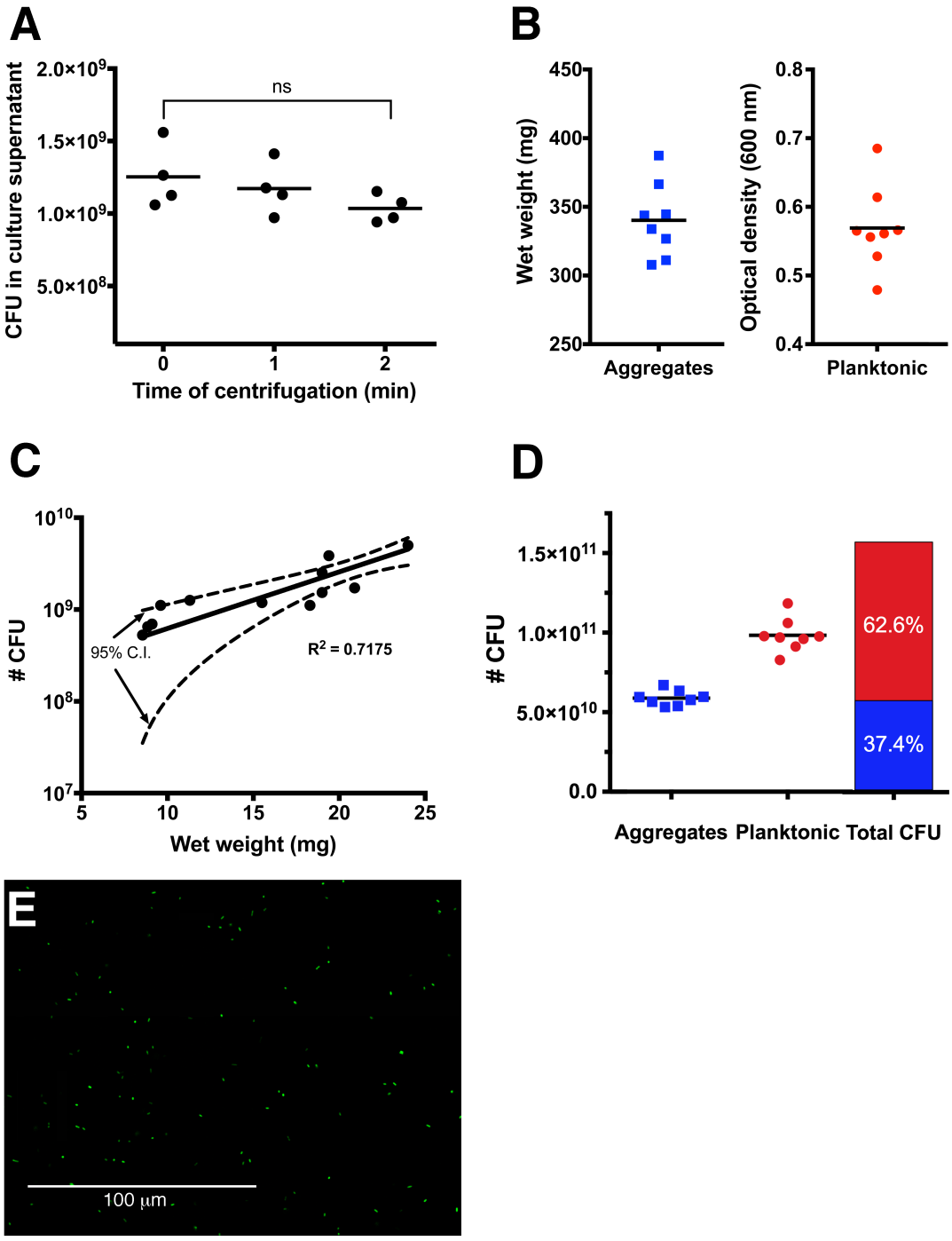

Figure 4.3. Isolation and quantitation of multicellular aggregates and planktonic cells and homogenization of multicellular aggregates.

(A) Culture aliquots were centrifuged at 210 x *g* for the times shown on the x-axis. The number of cells remaining in the supernatant was determined by performing serial dilutions followed by overnight growth on LB agar. The experiment was repeated three

times and the results of one representative experiment are shown. Statistical comparison between samples was performed using unpaired *t* tests. (B) Quantification of the two subpopulations present in replicate 100 mL cultures; for multicellular aggregates, the wet weight represents the total amount present in the culture flask, whereas for planktonic cells, the optical density is representative of the remaining culture media after removal of the aggregates. Horizontal bars represent the mean values. (C) Aggregates were resuspended in 1 mL of PBS and homogenized for 30 s using a tissue grinder, prior to performing serial dilutions. Each point represents the average from at least 3 technical replicates. The dotted lines represent the 95% confidence intervals (C.I.) of the data. (D) Colony forming units (CFU) present in aggregate or planktonic cell subpopulations from (B), as determined by the slope equation generated in (C) (aggregates) or by serial dilution plating (planktonic cells). The conversion factors were determined to be 1.73 x 10^8 CFU/mg of aggregates and 1.92 x 10^9 CFU/OD for planktonic cells. The bar graph represents the total CFU estimated within a biofilm flask culture and the relative proportion of cells within the aggregates (blue) and planktonic cell (red) subpopulations. The total number of planktonic cells within the culture is based on an approximated culture volume of 90 mL. (E) Aggregates from a *S.* Typhimurium culture containing a *csgBAC* promoter-GFP fusion plasmid were homogenized and visualized on a Leica TCS SP5 Confocal microscope. This image is representative of 60 slides that were prepared from three biological replicate cultures, showing that the homogenized suspension existed primarily of single cells. Small clusters of aggregated cells (i.e., <10 mm) were observed in ~3% of the fields of view analyzed (data not shown).

TP3) or undetectable levels (TP4) (Fig. 4.2B). We also analyzed both subpopulations for production of curli fimbriae, an important biofilm extracellular matrix component and functional bacterial amyloid (29). The major subunit protein of curli, CsgA, was detected in abundance in all aggregate fractions, whereas the planktonic cell fractions at TP2, TP3 and TP4 were negative (Fig. 4.2B). These results confirmed that the subpopulations had been cleanly separated by slow speed centrifugation. Trace amounts of CsgA were detected at TP1 (Fig. 4.2B), indicating that some cells had already begun to differentiate at this early time point.

The bacterial secondary messenger cyclic diguanylate monophosphate (c-di-GMP) is known to play a role in regulating *csgD* transcription and CsgD activity (30, 31). High levels of c-di-GMP are associated with biofilm formation whereas low levels are associated with a single cell lifestyle (44). To test if these same principles were at work in our system, we measured the intracellular concentration of c-di-GMP in planktonic cells at TP1 and in multicellular aggregates and planktonic cells at TP2. While planktonic cells at TP1 had nearly undetectable levels of c-di-GMP, there was an order of magnitude increase for planktonic cells at TP2 and another order of magnitude increase for aggregates, which had the highest levels (Fig. 4.2C). The levels of c-di-GMP correlated with the levels of CsgD in these cell types.

4.5.1 Transcriptome divergence between planktonic and aggregate cell types.

To get a comprehensive view of the transcriptional differences, RNA was purified from each cell subpopulation at each time point and directional RNA-seq was performed (Table 4.1). In total, 1856 genes or ~34% of genes in the genome were differentially expressed when comparing planktonic cells to multicellular aggregates at each time point (Tables A1 and A2; $p < 0.05$). The largest differences were identified at TP2, with 1240 genes being differentially expressed (Fig. 4.4A). At TP3 and TP4, the number of differentially expressed genes was reduced to 778 and 668, respectively. 642 of the differentially expressed genes were shared by at least two time points and 188 were shared for all three time points. From the 1856 genes that were differentially expressed, 784 were expressed at higher levels in multicellular aggregates (Fig. 4.4B) and 1072 were higher expressed in planktonic cells (Fig. 4.4C). We also performed a comparison

between the non-aggregative Δ*csgD* strain, which represents an entirely single cell population (13, 21), and the wildtype planktonic cells at TP2. Gene expression in planktonic cells from these two strains was more similar, with only 220 genes being differentially expressed (Tables A3 and A4).

Closer analysis of the activated genes revealed that the planktonic cells had higher expression of numerous pathways involved in virulence, including both *Salmonella* pathogenicity island (SPI) 1 and 2 T3SSs, part of the *phoPQ* regulon, motility and chemotaxis, and a high proportion (>40%) of function unknown (FUN) genes (Fig. 4.4D). In contrast, multicellular aggregates had higher expression of genes associated with extracellular matrix production in biofilms (i.e., *csgDEFG*, *csgBAC* and *yaiC* (*adrA*)), changes in metabolism (i.e., amino acid, carbohydrates, lipids and nucleotides), and pathways associated with stress response (i.e., osmoprotection) (Fig. 4.4D). The nature of the transcriptional differences between planktonic cells and multicellular aggregates suggested that they represented specialized cell types. Many of the genes listed in Datasets 4.1 and 4.2 were distributed into operons predicted from RNA-seq through a combination of transcriptional start and termination sites, intergenic distance and co-expression information (Y. Wang, K.D. MacKenzie and A.P. White, submitted for publication).

4.5.2 Expression of c-di-GMP enzymes and other genetic factors contributing to CsgD bistability.

We identified several differentially expressed genes associated with the complex c-di-GMP signaling system that contributes to CsgD regulation (Table 4.2). Two diguanylate cyclases (i.e., c-di-GMP-generating enzymes) *STM1987* (45) and *adrA* (46), were more highly expressed in aggregates at TP2, which correlated with higher levels of c-di-GMP in these cells. *yhjH*, which encodes a phosphodiesterase (i.e., c-di-GMP-degrading enzyme) with a critical role in controlling motility (46, 47), was more highly expressed in planktonic cells, and this correlated with lower amounts of c-di-GMP. Three diguanylate cyclases characterized as having positive effects on invasion (*yeaJ*; (48)), motility (STM2503; (49)) or motility and virulence (STM4551; (49)), were more highly

Table 4.1. RNAseq-based transcriptome analysis of planktonic cells and multicellular aggregates isolated from *S.* Typhimurium culture.

	Wildtype planktonic				Wildtype aggregates			Δ*csgD* planktonic
Characteristic	**TP1**	**TP2**	**TP3**	**TP4**	**TP2**	**TP3**	**TP4**	**TP2**
Total no. of reads[a]	50,889,232	64,194,848	60,641,580	66,424,412	64,372,950	51,211,612	59,213,696	32,648,668
mRNA & sRNA	5,142,914	9,183,473	6,970,035	4,208,777	7,112,841	4,395,392	4,089,516	3,971,434
Genome Coverage[b]	77.7	138.7	105.3	63.6	107.5	66.4	61.8	61.2
Protein-coding genes[c]								
Sense (%)	75.3	74.0	68.9	66.0	74.8	71.6	70.3	73.9
Antisense (%)	5.9	5.4	6.7	8.9	2.8	4.1	4.4	10.4
sRNAs (%)[d]	8.0	10.6	14.2	9.5	14.4	13.9	8.9	6.5
Intergenic regions (%)	10.8	10.0	10.2	15.6	8.0	10.4	16.4	9.1
Plasmid (%)	0.8	0.8	0.7	0.6	0.4	0.4	0.4	0.8

[a] Reads of three technical replicates were merged for each time point of each sample. In total, 416 million strand-specific reads of 75 nt length were obtained.

[b] The reads mapped to rRNAs and tRNAs were removed prior to statistical and comparative analysis. This yielded an average genome coverage of 89.9 × (±23.2) for each time point of each sample excluded.

[c] There are 5,414 putative ORFs in the *S.* Typhimurium ATCC 14028 genome, including 5,313 chromosome-encoding and 101 plasmid-encoding ORFs.

[d] The full set of known sRNAs in *Salmonella* was annotated from (40). Approximately 11.5 ± 2.8% of the mapped reads encoded known sRNAs and 11.8 ± 3.4% of the reads mapped to additional intergenic regions.

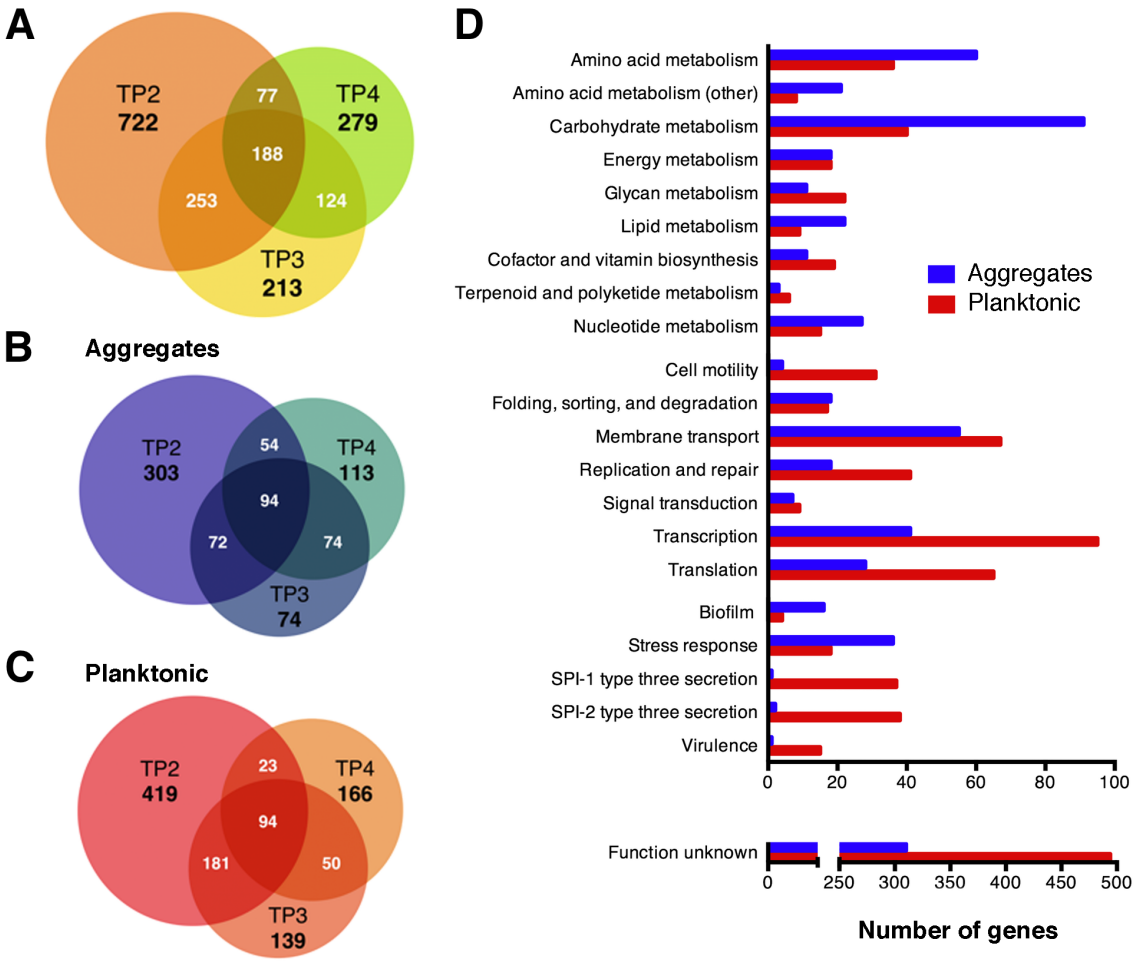

Figure 4.4. RNAseq-based transcriptome comparison between multicellular aggregate and planktonic cell subpopulations.

(A) The Venn diagram shows the total number of differentially expressed genes (p <0.05) between planktonic cells and multicellular aggregates at each time point. This total was subdivided into genes that were higher expressed in aggregates (B), and genes that were higher expressed in planktonic cells (C). (D) Differentially expressed genes in aggregates versus planktonic cells were categorized using the KEGG pathway database (Kyoto Encyclopedia of Genes and Genomes; http://www.genome.jp/kegg/) and are listed in Dataset S1.

expressed in planktonic cells at TP2, which correlated with the predicted phenotype of these cells. *nlpI*, encoding a putative membrane protein with a repressive effect on CsgD synthesis (50) was also more highly expressed in planktonic cells, which matched the lower levels of CsgD in these cells. In contrast, high expression of *ycgR* and *ydiV* in planktonic cells at TP2 did not match the predicted phenotypes. YcgR contains a PilZ domain for binding c-di-GMP and represses motility by acting as a type of flagellar brake (51, 52). Based on the published function, we would have expected *ycgR* to be more highly expressed in multicellular aggregates and to function at higher c-di-GMP levels inside the cell. YdiV represses expression of *yhjH* and also negatively regulates motility (53). It is possible that the synthesis of YcgR and YdiV result in localized effects within the cell, but the overall c-di-GMP pool is predominantly influenced by the increased expression of *yhjH* (46, 47). In general, the RNA-seq findings were consistent with the levels of c-di-GMP detected and the predicted phenotypes of planktonic cells and multicellular aggregates.

Bistable gene expression in bacteria is usually the result of positive autoregulation of a master regulatory gene (54, 55). There has been a feed-forward loop proposed for *csgD* in *E. coli*, but it has not been proven in *S.* Typhimurium (30). CsgD is proposed to activate expression and synthesis of IraP, which inhibits the degradation of RpoS, thus leading to greater *csgD* transcription (56). RNAseq analysis revealed that multicellular aggregates had elevated transcription of *rpoS* at TP2 and *csgD* and *iraP* (*yaiB*) at TP3 and TP4, potentially confirming this regulatory loop (Fig. 4.5).

4.5.3 SPI-1 T3SS expression and synthesis and motility are increased in planktonic cells

The SPI-1 T3SS is a 3.5 megadalton translocation apparatus that is required for injection of *Salmonella* effector proteins into the cytoplasm of host cells. These effector proteins stimulate *Salmonella* invasion and dissemination, as well as host inflammation (6). The current dogma is that SPI-1 T3SS synthesis is activated by the body temperatures of warm-blooded hosts (57). This is primarily based on the finding that the nucleoid-associated protein H-NS represses transcription of key SPI-1 regulators at lower growth temperatures, such as 25°C (58). Biofilm growth conditions (i.e., low osmolarity, 28°C, stationary phase) were traditionally thought to be non-SPI-1 inducing because they

Table 4.2. Differentially expressed genes in the c-di-GMP signaling network.

Gene Name[a]	Enzyme type[b]	Function[c]	Fold change[d]	Time points[e]
STM14_5467 (STM4551)	DGC	Positive regulator of motility and virulence by both c-di-GMP dependent and independent mechanisms (49); an additional report states that function is the exact opposite (48)	4.05	TP2|TP3|TP4
yeaJ (STM1283)	DGC	Stimulates invasion (48)	3.18	TP2|TP3
yhjH (STM3611)	PDE	Negative regulator of rdar morphotype; positive regulator of motility (46)	4.47	TP2
STM14_3068 (STM2503)	DGC	Positive effect on motility (49)	3.13	TP2
ycgR	c-di-GMP biosensor	Binds c-di-GMP through a high affinity PilZ domain; represses motility by binding directly to FliG (51, 52)	4.07	TP2
ydiV (STM1344)	EAL-like protein	Positive regulator of *csgD* transcription and negative regulator of motility, but does not have a direct role in c-di-GMP turnover; represses expression of *yhjH (53)*	3.58	TP2
nlpI	Outer membrane protein	NlpI can negatively influence the expression of *yjcC*, which codes for a DGC affecting CsgD expression (50)	3.89	TP2|TP3|TP4
adrA (*yaiC*)	DGC	Positive regulator of cellulose production (46)	0.34	TP2|TP3|TP4
STM14_2408 (STM1987)	DGC	Positive effect on rdar morphotype, CsgD expression, cellulose (45)	0.54	TP2

[a] Gene names are listed as they appear in the *S.* Typhimurium 14028 genome (39); the *S.* Typhimurium LT2 gene names are listed in parenthesis.
[b] An enzyme that generates c-di-GMP is called a diguanylate cyclase (DGC), an enzyme that breaks down c-di-GMP is called a phosphodiesterase (PDE). PDE enzymes contain EAL domains.
[c] Functions are listed in the context of c-di-GMP regulation and turnover, CsgD expression, motility, and virulence.
[d] Fold-change represents the expression ratios of planktonic / aggregates at TP2; genes with values >1 have higher expression in planktonic cells, whereas genes with values <1 have higher expression in multicellular aggregates.
[e] Time points are listed where the expression difference between planktonic cells and multicellular aggregates was statistically significantly (p <0.05).aggregates had elevated transcription of *rpoS* at TP2 and *csgD* and *iraP* (*yaiB*) at TP3 and TP4, potentially confirming this regulatory loop (Fig. 4.5).

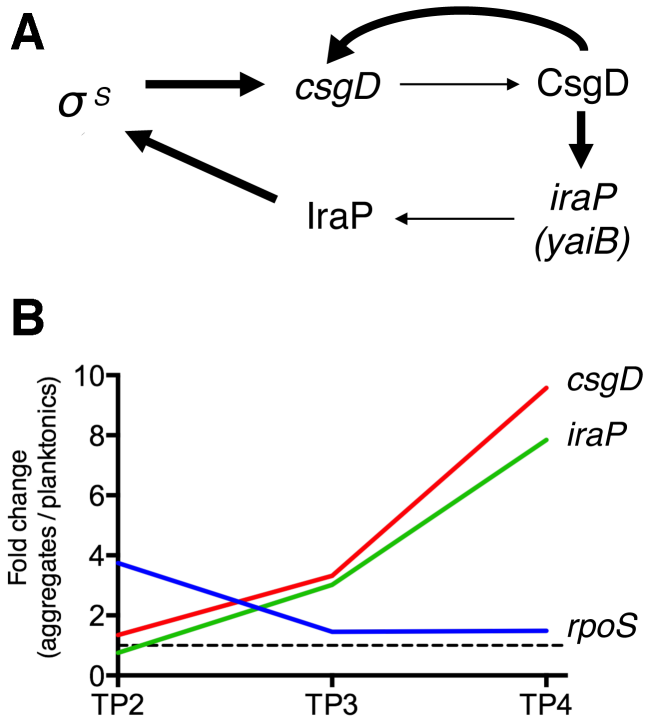

Figure 4.5. The feed forward regulatory loop for CsgD activation.

(A) The regulatory network representing a potential feed-forward loop for CsgD activation (Gualdi, L., L. Tagliabue, and P. Landini. 2007. J. Bacteriol. 189: 8034-8043). σ^S represents the RpoS sigma factor, which is required for activation of *csgD* transcription. Thick black lines represent regulation of transcription, with the exception that IraP stabilizes RpoS protein; thin black lines represent protein synthesis. We added a thick line representing auto-regulation of CsgD (Ogasawara, H., K. Yamamoto, and A. Ishihama. 2011. J Bacteriol. 193: 2587-97). (B) Relative mRNA expression for each of the genes shown at TP2, TP3 and TP4. The dotted line represents no difference in gene expression between multicellular aggregates and planktonic cells.

are so different from classical SPI-1 inducing conditions (i.e., high osmolarity, 37°C, late-exponential phase) (57, 59, 60).

RNA-seq analysis revealed that 41 of 44 SPI-1 T3SS related genes were more highly expressed in planktonic cells grown under biofilm-inducing conditions (Dataset 4.1). These genes included pivotal transcriptional regulators (Fig. 4.6A), as well as apparatus, chaperone and effector genes (Fig. 4.6B). HilA activates transcription of SPI-1 apparatus genes distributed in the *inv/spa* and *prg/org* operons (57). InvF is encoded as part of the *inv/spa* operon, and together with SicA promotes transcription of the *sic/sip* operon encoding the translocon, as well as numerous effector genes. Temporal expression analysis of the SPI-1 genes followed a pattern consistent with the establishment of secretion-competent organelles in planktonic cells; apparatus genes had highest expression at TP1, while translocon and numerous effector genes had peak expression at TP2 (Fig. 4.7). In contrast, expression levels of all SPI-1 genes were reduced in multicellular aggregates (Fig. 4.7). Comparison between wildtype planktonic cells and Δ*csgD* planktonic cells at TP2 revealed that *sipD* and *sopE2* were higher expressed in the wildtype cells (Dataset 4.2). In addition, several other SPI-1 genes were higher expressed in wildtype cells but had *p* values just above the cut off for significance (data not shown). These findings suggested that there could be a potential regulatory link between CsgD and the SPI-1 T3SS.

To determine if SPI-1 transcriptional differences were reflected at the protein level, whole cell lysates of either cell type were analyzed by western blotting. InvG, the OM secretin (61), was detected at high levels in planktonic cells at all time points and was nearly undetectable in the multicellular aggregate samples (Fig. 4.6C). The on/off pattern of synthesis for this critical structural component indicated that planktonic cells had the capacity to form a greater number of SPI-1 T3SS organelles than aggregated cells. SipD, the translocon tip protein that interacts with host cell membranes (62), was also present at high levels in planktonic cells at all time points (Fig. 4.6C). SopB and SopE2, two of the earliest expressed effectors that stimulate host cell cytoskeleton rearrangements (63), were also synthesized at higher levels in planktonic cells at TP2 and

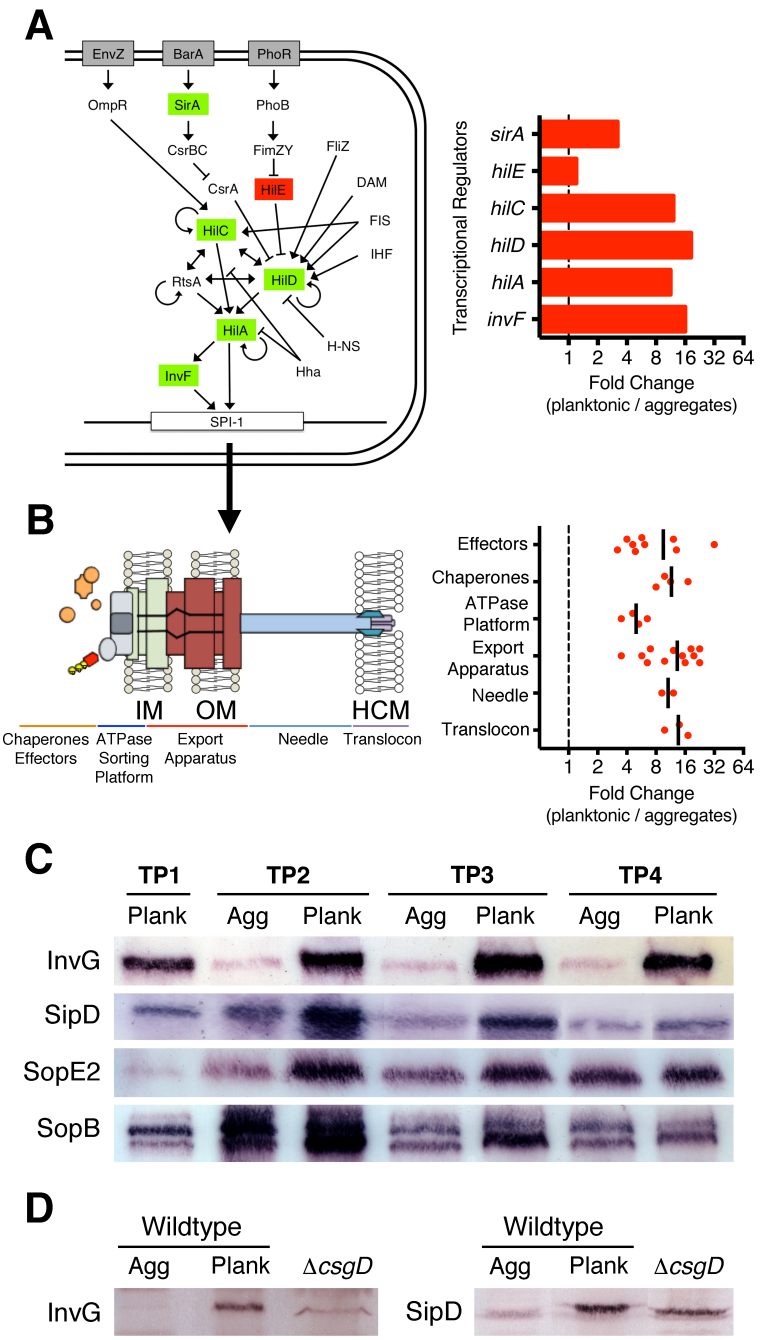

Figure 4.6. Expression and synthesis of the SPI-1 T3SS in *S.* Typhimurium planktonic cells.

(A) A schematic of a bacterial cell showing the regulatory map and main transcriptional activators (green) or repressors (red) for SPI-1 T3SS expression, adapted from (6). The fold-change values for the regulators shown are displayed as red bars on the accompanying graph, calculated from RNA-seq data at TP2. (B) The SPI-1 T3SS

translocation apparatus (left), adapted from (90), was divided into five component groups, including proteins embedded in the bacterial inner or outer membrane or host cell membrane (IM, OM, HCM). Each point on the accompanying graph represents the fold change value for the following genes at TP2 for Effectors: *sopE2, sopB, sipA, sopD, gtgE, sopA, avrA, sptP, gogB*; Chaperones: *sicA, sicP, invB, sigE (pipC)*; ATPase platform: *invC, spaO, orgA, orgB*; Export apparatus: *invG, spaR, prgH, invI*, *invA*, *invE*, *prgK*, *spaP*, *spaQ*, *invJ*, *spaS*, *invH*; Needle, *prgI, prgJ*; and Translocon: *sipB*, *sipC*, *sipD* (6). The dotted line at 1 represents no difference in gene expression; the vertical lines represent the mean fold-change for the group. (C) Whole cell lysates of planktonic cells and multicellular aggregates were analyzed by western blotting to detect the SPI-1 T3SS proteins as shown; Plank = planktonic cells, Agg = multicellular aggregates. (D) InvG and SipD were detected in Δ*csgD* planktonic cells at TP2 and the relative levels were compared to aggregates and planktonic cells isolated from wildtype cultures.

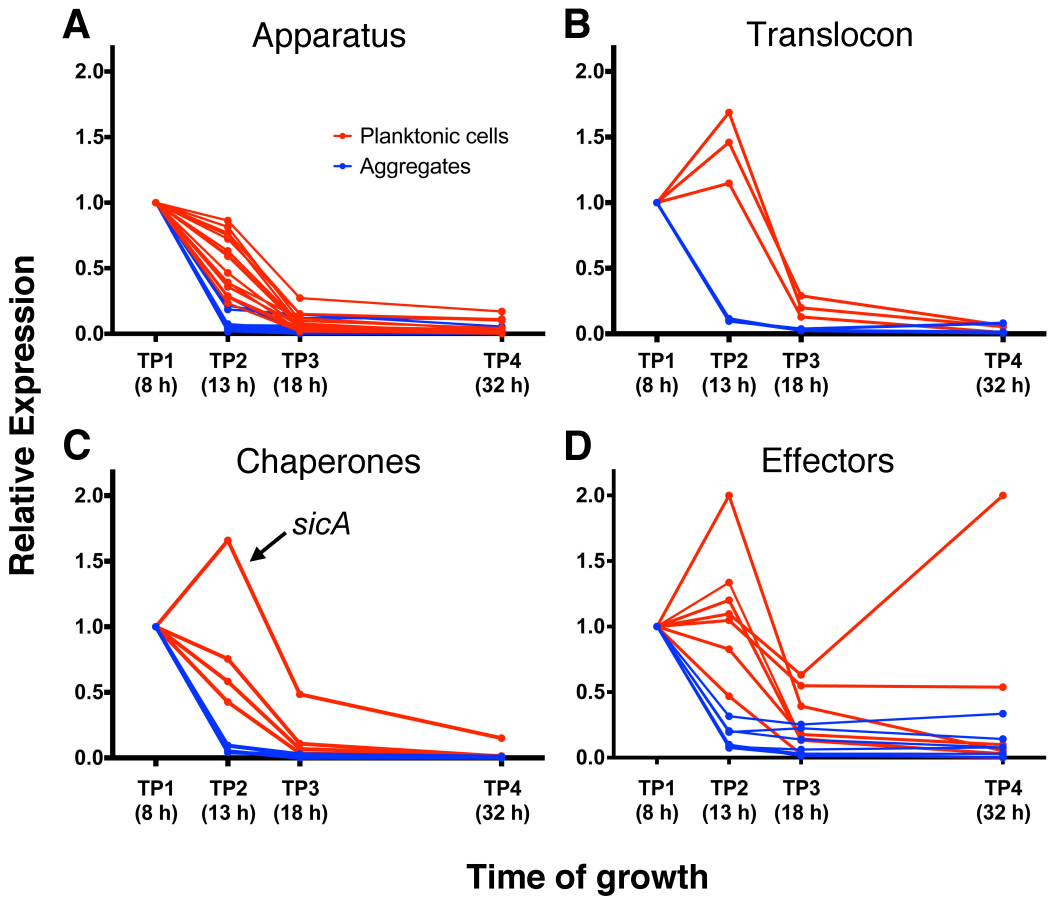

Figure 4.7. Transcriptional dynamics of genes associated with the SPI-1 type three secretion system.

The relative mRNA expression of SPI-1 T3SS apparatus (A), translocon (B), chaperone (C), and effector (D) genes in each subpopulation of cells was compared across time points. Genes were assigned to the same functional classes as listed in Figure 4B. Relative expression was obtained by dividing the absolute expression level (i.e., number of mapped cDNA reads obtained by RNA-seq) for each gene at each time point by the absolute expression level of that gene at TP1. The starting value for each gene is at 1.0. The expression profile of *avrA* did not match the other effector genes; therefore it was not included in (D).

TP3, with the difference decreasing by TP4 (Fig. 4.6C). SPI-1 protein levels were also analyzed in Δ*csgD* cells at TP2. InvG and SipD levels were noticeably lower in Δ*csgD* cells as compared to wildtype planktonic cells (Fig. 4.6D), whereas SopB and SopE2 had no detectable difference between the two strains (data not shown).

Recent evidence suggests that the SPI-1 T3SS and flagellar-mediated motility are required for targeting of permissive invasion sites on host cells (64) and these systems are often co-regulated (65). In our RNA-seq experiment, 28 genes coding for flagella assembly and chemotaxis proteins were identified as being more highly expressed in planktonic cells (Dataset 4.1). We performed a swim assay, which confirmed that planktonic cells had enhanced motility compared to multicellular aggregates, but were not significantly different from Δ*csgD* planktonic cells (Fig 4.8).

4.5.4 Transcriptional priming of the SPI-2 T3SS in planktonic cells

Salmonella pathogenicity island 2 (SPI-2) encodes a second T3SS that has an important role once *Salmonella* has invaded host cells. The SPI-2 T3SS is required for modification of the phagolysosome to make it conducive for *Salmonella* replication, particularly inside of macrophages (6). 37 of 42 genes encoding for the SPI-2 T3SS were more highly expressed in planktonic cells as compared to the multicellular aggregates (Dataset 4.1). However, unlike SPI-1, we were unable to detect SPI-2 protein synthesis, such as SsaC, the equivalent of InvG, in any cell fractions (data not shown). This suggested that functional SPI-2 needles were not present on the surface of planktonic cells and supported the possibility that transcriptional priming was occurring, as has been observed prior to host cell invasion (66, 67).

4.5.5 Planktonic cells display enhanced invasion of a human intestinal cell line

To our knowledge, SPI-1 T3SS synthesis has never been observed before under biofilm-inducing conditions, but it is usually correlated with increased invasion and virulence (68). Therefore, we tested the invasion of different *S.* Typhimurium cell types into polarized Caco-2 cells, a human intestinal cell line. Planktonic cells isolated at TP2 displayed the same invasiveness as planktonic cells grown under classical SPI-1 inducing conditions (Fig. 4.9), confirming that the SPI-1 organelles synthesized under biofilm-

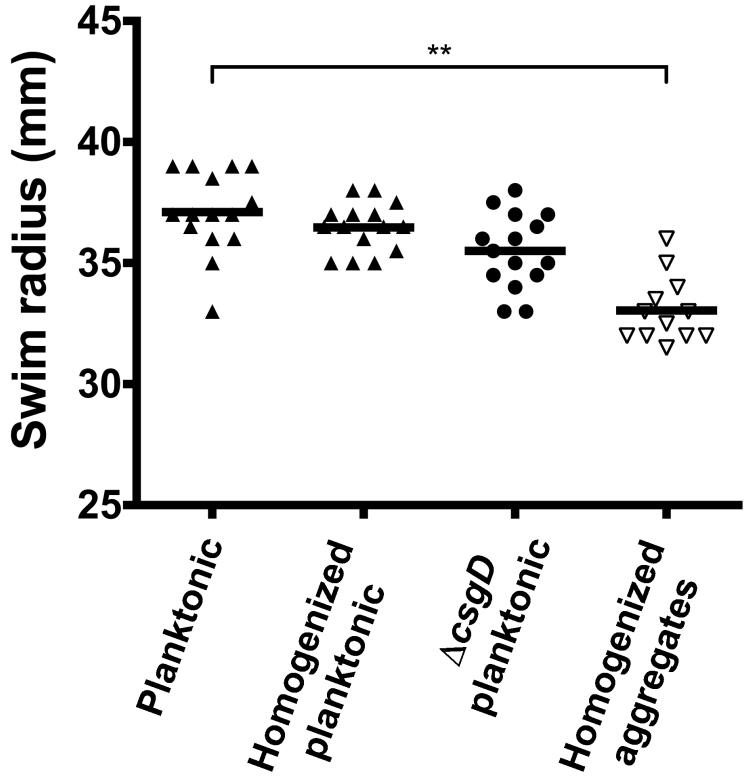

Figure 4.8. Swim assay of multicellular aggregates and planktonic cells.

Cell subpopulations were isolated from *S. Typhimurium* biofilm cultures at TP2, and either tested for motility directly or after homogenization for 30s in a tissue grinder. Swim plates (lysogeny broth (Luria) + 0.25% agar) were inoculated with 5 x 10^6 CFU and incubated at 37°C for 8 h prior to measuring the swim radius. Statistical comparison between samples was performed using an unpaired *t* test (**, $p < 0.01$).

inducing conditions were functional. Planktonic cells isolated at later time points (i.e., TP4) still retained efficient invasion, suggesting that this phenotype was stable. In contrast, the multicellular aggregates were significantly less invasive, presumably because of reduced type three secretion synthesis. Surprisingly, Δ*csgD* cells, which had reduced InvG and SipD levels, only showed a slight drop in invasion that was not statistically different than wildtype planktonic cells. For the aggregates, we reasoned that reduced invasion could be influenced by the presence of extracellular matrix polymers blocking access to the Caco-2 cell surface. To disrupt the matrix, we homogenized the aggregates prior to repeating the assay, but it did not lead to increased invasion (Fig. 4.9, homogenized aggregates). It was noted, however, that the aggregated cells invaded significantly better than a ΔSPI-1 deletion strain, suggesting that they would retain a capacity for virulence.

As part of the Caco-2 cell invasion assays, the aminoglycoside antibiotic gentamicin was added to kill any extracellular bacteria that had not invaded. Prior to performing the assays, we tested the susceptibility of multicellular aggregates and planktonic cells to gentamicin. While the planktonic cells were completely killed (below the limit of detection) at standard concentrations of 200 – 400 μg mL^{-1} (33) the multicellular aggregates displayed enhanced resistance and were only depleted once the concentration reached 1200 μg mL^{-1} (Fig. 4.10). This confirmed that the multicellular aggregates represented a more resistant subpopulation of cells.

4.5.6 Planktonic cells display a significant virulence advantage in murine infection.

Based on the expression data and Caco-2 invasion results, we hypothesized that the planktonic cells were more virulent, whereas the multicellular aggregates were likely to have enhanced persistence. To test the colonization efficiency and virulence of each cell type, we performed formally randomized, blinded co-infections of mice with 1:1 ratios of multicellular aggregates and planktonic cells isolated from kanamycin- or chloramphenicol-resistant strains of *S.* Typhimurium (Fig. 4.11). For the initial competitive index (CI) trial, we wanted to keep the cell subpopulations in their native physical state; therefore, we minimized the handling of cells and used a wide bore gavage

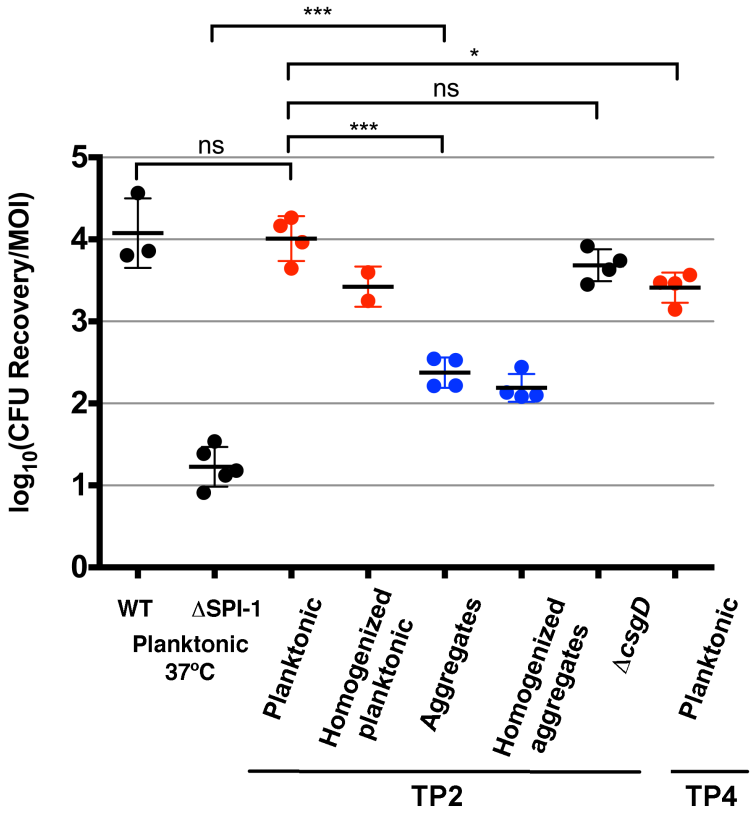

Figure 4.9. Invasion of polarized Caco-2 cells by different *S.* Typhimurium subpopulations.

Planktonic cells and multicellular aggregates isolated from *S.* Typhimurium 14028 biofilm cultures at TP2 and TP4 were used to infect polarized Caco-2 cells. Aggregates that were homogenized using a tissue grinder were also tested, along with planktonic cells prepared from Δ*csgD* cultures. Wildtype and ΔSPI-1 cells grown in LB at 37°C were used as positive and negative invasion controls, respectively. Each point represents the average value from invasion of three replicate wells of Caco-2 cells; the horizontal bars represent the mean values. The average MOI was 17, based on an estimate of 2.5 x 10^5 polarized Caco-2 cells per well. Statistical significance is noted as: ***, p <0.001, *, p <0.05, ns, p >0.05.

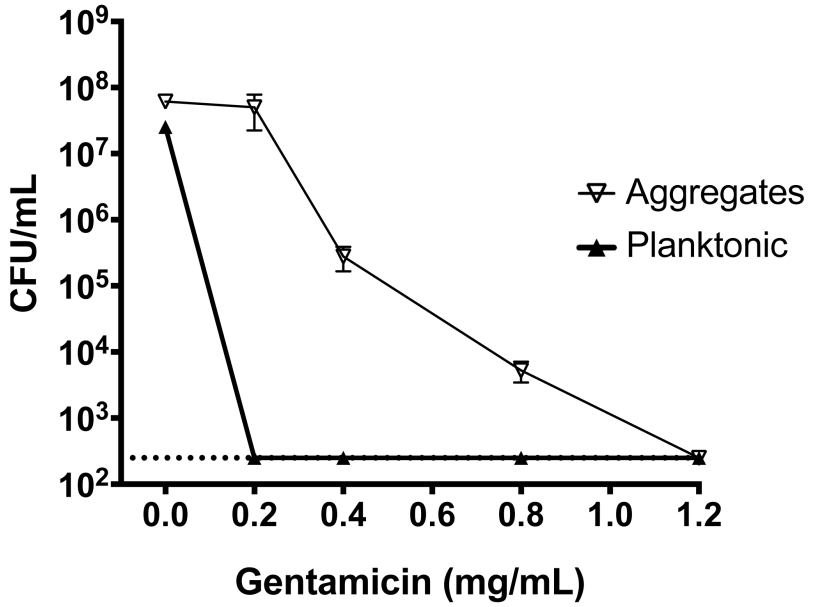

Figure 4.10. Antimicrobial killing assay of multicellular aggregates and planktonic cells exposed to gentamycin.

Multicellular aggregates and planktonic cells were exposed to gentamycin for 2 h at the concentrations shown on the x-axis. The number of surviving cells is plotted on the y-axis; the dotted line represents the limit of detection of 250 CFU. Points and error bars represent the means and standard deviations from three biological replicate experiments. Statistical differences in viable cell counts between the cell types was performed using an unpaired *t* test; *, $P < 0.05$, ****, $P < 0.0001$.

needle to prevent shearing of aggregates during oral delivery. Under these conditions, the planktonic cells exhibited a significant virulence advantage, as demonstrated by their proportionally increased recovery from spleen, liver, cecum and mesenteric lymph nodes in the majority of infected mice (Fig. 4.12A, red triangles). As predicted, the multicellular aggregates did retain a capacity for virulence and were recovered at higher proportions in several mice (Fig. 4.12A, blue triangles). We reasoned that despite being comprised of hundreds to thousands of cells, each aggregate might only function as a single infectious unit. Therefore, we homogenized the aggregates to generate a single cell suspension prior to repeating the CI trial. The CI values were slightly reduced, but the overall trend remained the same (Fig. 4.12B), indicating that the increased virulence of planktonic cells was not due to a difference in multiplicity of infection. We also competed wildtype planktonic cells and Δ*csgD* planktonic cells, and again the wildtype planktonic cells displayed a strong virulence advantage (Fig. 4.12C). This was surprising because the invasion phenotypes of these two strains were similar; however, there could be other impairments related to the Δ*csgD* mutation (69). To assess the importance of SPI-1 in the colonization advantage displayed by planktonic cells, we repeated a CI trial with planktonic cells and multicellular aggregates prepared from a ΔSPI-1 mutant strain. For each organ, the median CI values dropped approximately 10-fold and were near 1 (Fig. 12D), which indicated that the SPI-1 T3SS was the primary factor responsible for increased colonization of planktonic cells. However, small differences were still detected for the spleen, liver and cecum (Fig. 4.12D), indicating that the planktonic cells had other adaptations (e.g., increased motility) that provide them with a colonization advantage.

4.5.7 Persistence of multicellular aggregates reduces the competitive advantage of planktonic cells

To assess how the multicellular aggregates and planktonic cells responded to a potential environmental stress, aliquots of the challenge inocula from the first CI experiment were desiccated and stored in 24-well plates at room temperature. Viability assays performed on the desiccated challenge inocula revealed that the multicellular aggregates survived significantly better than the planktonic cells (Fig. 4.13A). At the

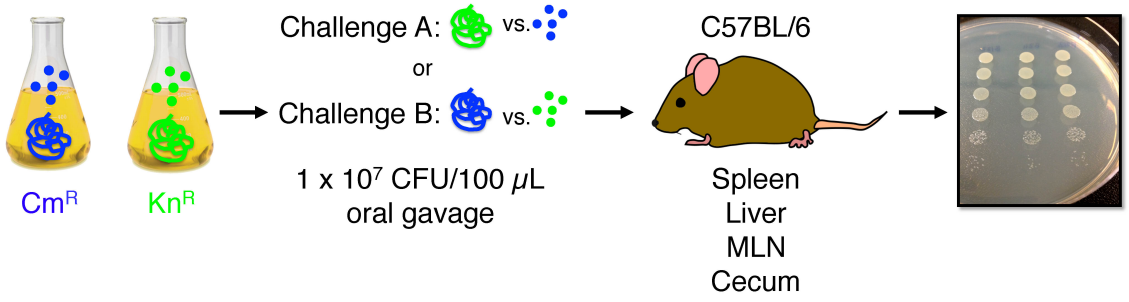

Figure 4.11. Preparation of mixed bacterial challenges of multicellular aggregates and planktonic cells for competitive infection assays.

Cell subpopulations were isolated from Cm^R or Kn^R *S.* Typhimurium cultures and used for reciprocal co-infections of C57BL/6 mice. We aimed for a challenge dose of 10^7 CFU consisting of an approximate 1:1 ratio of planktonic cells and multicellular aggregates. Mice were randomly assigned into four groups of six mice; challenge A and challenge B was given to two groups each. At 4-7 days post infection, mice were euthanized and their organs were homogenized to determine the number of CFU of Cm^R or Kn^R *S.* Typhimurium that were present; MLN, mesenteric lymph nodes.

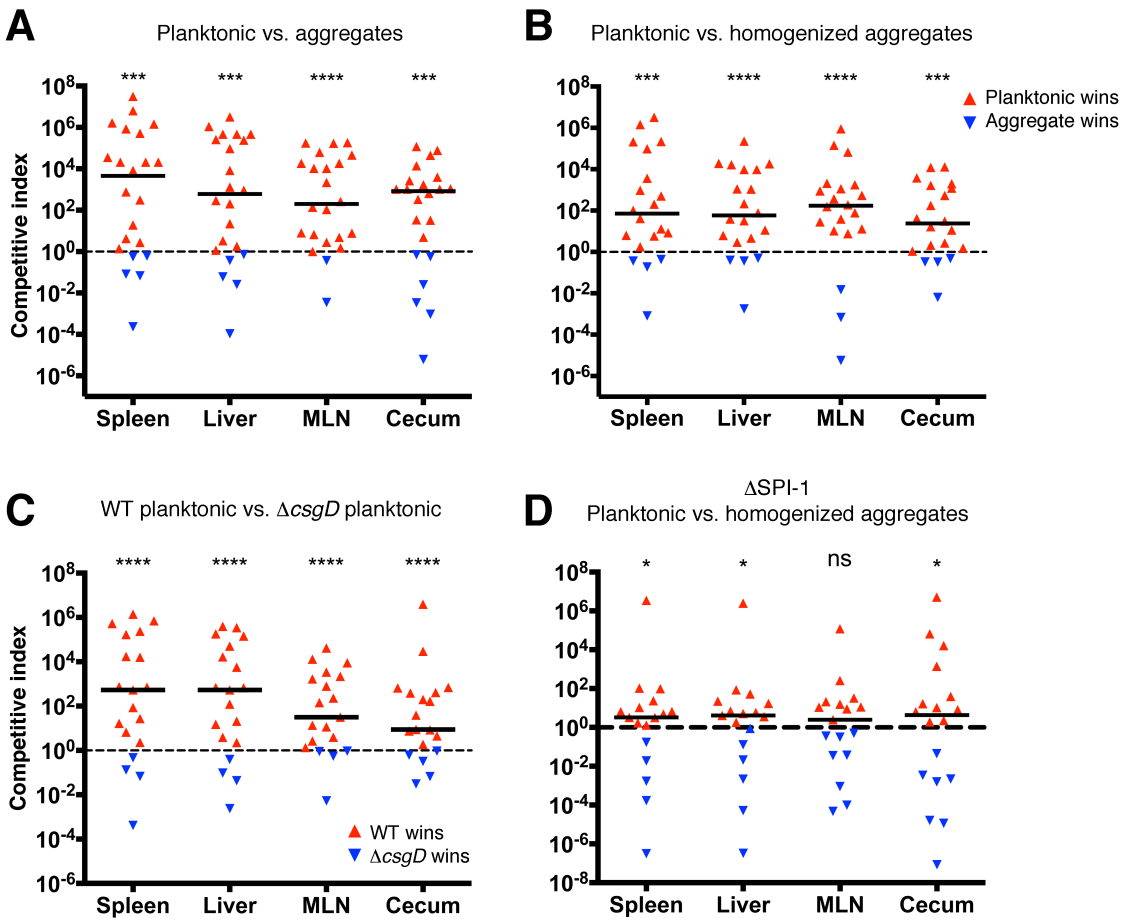

Figure 4.12. Competitive infections between multicellular aggregate and planktonic cells prepared from wildtype, ΔSPI-1 and Δ*csgD* cultures.

The competitive index (CI) was determined by oral co-infection of C57BL/6 mice with a 1:1 ratio of different cell types. (A) Trial between planktonic cells and multicellular aggregates, using a wide-bore gavage needle to prevent shearing of aggregates. (B) Trial between planktonic cells and multicellular aggregates after homogenization of the aggregates to generate a single cell population. (C) Trial between wildtype planktonic cells and Δ*csgD* planktonic cells. (D) Trial between planktonic cells and multicellular aggregates isolated from a ΔSPI-1 biofilm culture. Each point in A-D represents the CI value calculated from one organ from a single mouse, with either cell type winning the competition (red or blue triangles). The solid lines represent the median of all CI values from a particular organ. The dotted line represents a CI value of 1, which indicates no difference in virulence. MLN = mesenteric lymph nodes. Statistical significance is noted as: * p <0.05, ** p <0.01, *** p<0.001, **** p <0.0001, ns p >0.05.

four-week time point, there were approximately 10 times more aggregated cells alive (Fig. 4.13A, dotted line). We performed another CI trial at this time point, after rehydration and resuspension of the cells. The values for each organ indicated no significant advantage for either subpopulation, with aggregates or planktonic cells winning the competition in an approximately equal number of mice (Fig. 4.13B). This indicated that upon exposure to environmental stress, the multicellular aggregates had a fitness advantage. If these results were extrapolated into the natural world, one could envision a scenario where after extended periods of environmental exposure only multicellular aggregates would be alive and be able to infect new hosts.

4.6 Discussion

In this study, we isolated and analyzed two distinct subpopulations of cells within a clonal population of *S.* Typhimurium. Based on genetic, biochemical and functional characterization, we hypothesize that the planktonic cells and multicellular aggregates formed under biofilm-inducing conditions represent specialized cell types that are adapted for virulence and persistence, respectively. Due to the presence of functional SPI-1 T3SS organelles, we predict that the planktonic cells are proficient at direct host-to-host transmission, whereas the resistant, multicellular aggregates can survive in the environment and cause infections at later times. The connection of virulence and persistence within an individual population represents an elegantly simple strategy that would prepare *S.* Typhimurium for unpredictability and improve the chances for transmission each time they leave an infected host. This scenario also implies that the host immune system may respond differently to these specialized cell types. If hosts routinely encounter multicellular aggregates, it could explain why curli fimbriae are such potent stimulators of the innate immune system (29, 70, 71) even though these organelles do not appear to play a role in *Salmonella* infections (25).

The concept of bistability was described by Novick and Weiner over 50 years ago as the existence of different subpopulations of enzyme-producing cells within a genetically identical population of *E. coli*, controlled by a feed forward loop (72). CsgD lies at the heart of a feed-forward loop in *S.* Typhimurium when cells are grown under biofilm-inducing conditions. CsgD is synthesized at high levels in specific cells, leading

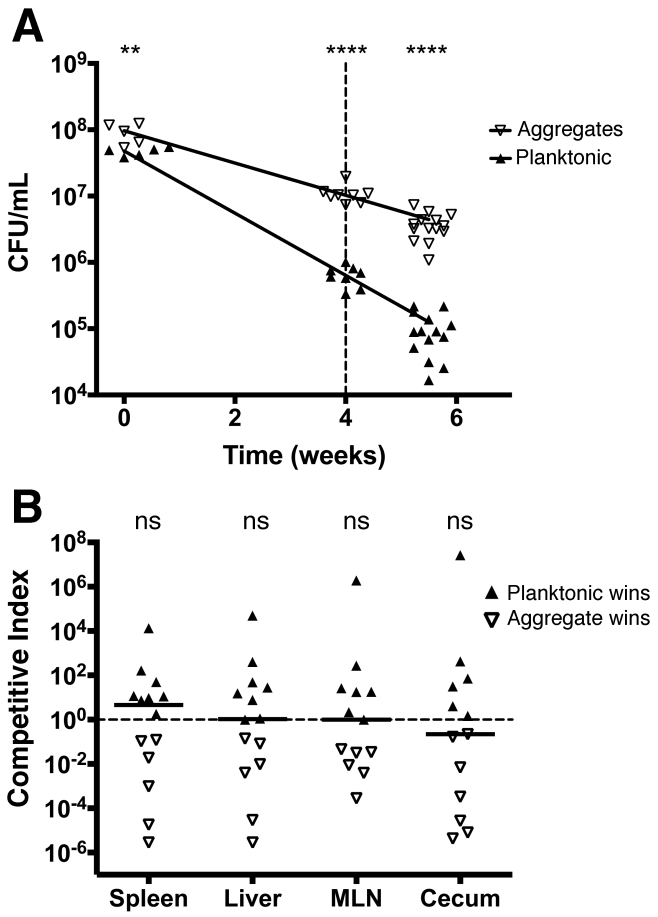

Figure 4.13. Enhanced persistence of *S.* Typhimurium multicellular aggregates.

(A) Multicellular aggregates and planktonic cells used for the trial in Fig. 6A were desiccated and stored in 24 well plates. Survival of each cell type was determined after storage for the times displayed on the x-axis. Statistical significance is noted as: * p <0.05, ** p <0.01, **** p <0.0001. The dotted line represents the time at which cells were resuspended and a CI trial was performed (B). Each point represents the CI value calculated from one organ from a single mouse and the solid lines represent the median of all CI values from a particular organ. The dotted line represents a CI value of 1, or no difference in virulence; none of the organs displayed CI values that were significantly different from 1 (ns = p > 0.05).

to production of extracellular matrix polymers and the formation of multicellular aggregates (25, 30). The regulation of *csgD* transcription and CsgD synthesis is extremely complex (reviewed in (31)). Numerous transcription factors are involved, including OmpR, MlrA, and CpxR, non-specific DNA-binding proteins H-NS and IHF, in addition to the feed-forward loop involving *rpoS* and the factors that regulate RpoS activity, including IraP and Crl. In addition, at least five sRNAs can bind to the 5' untranslated region of the csgD mRNA and regulate CsgD synthesis (73). Robust control of CsgD production is clearly important, given that > 30% of genes in the genome were differentially expressed between the *S.* Typhimurium multicellular aggregates and planktonic cells. The large number of differentially expressed genes represents a tremendous amount of non-heritable genetic change, on the order of a developmental program in bacteria (74). Since only 20 CsgD-specific targets were identified in a recent *E. coli* ChIP-seq study (75), we hypothesize that CsgD initiates the aggregation process, and a cascade of gene expression changes follows as cells produce an extracellular matrix (35). Ultimately, this cascade would be the cumulative result of numerous connected signaling pathways, including c-di-GMP (44).

Recent reports have shown that bistable gene expression can occur within bacterial virulence gene networks. In *Vibrio cholerae*, bimodal expression of ToxT results in virulent and non-virulent cell subpopulations (76). Bistable production of the SPI-1 T3SS in *S.* Typhimurium has also recently been described (77-79). These authors observed that nearly 100% of *S.* Typhimurium cells interacting with intestinal tissues were positive for SPI-1 expression while only 15% of cells in the intestinal lumen were positive (77). The presence of both SPI-1^+ and SPI-1^- cells seems important for stabilizing the virulence traits of *S.* Typhimurium (78). At the moment, it is unclear how these findings are connected to what we have observed. Our finding that SPI-1 T3SS synthesis occurs in planktonic cells under biofilm-inducing conditions may indicate that an alternative regulatory pathway exists for SPI-1 activation. It is also possible that the planktonic population consists of both SPI-1^+ and SPI-1^- cells. Alternatively, the subpopulation of cells that form multicellular aggregates under biofilm-inducing conditions could be related to the SPI-1^- cells described by WD Hardt and colleagues (e.g., similar genetic pathways activated or repressed). One of our initial goals was to

determine if the differentiation of cells was more than an *in vitro* phenomenon. Since the planktonic cells were more virulent than multicellular aggregates, even after the aggregates were homogenized into a single cell suspension, it suggested that differentiation was retained until cells reached the murine intestine. The importance of SPI-1 in this process was confirmed since planktonic cells and multicellular aggregates from a ΔSPI-1 mutant strain had nearly equal colonization efficiencies. Furthermore, the reduction in virulence that we observed for the Δ*csgD* strain, together with recent connections between c-di-GMP, virulence (28) and motility (52) suggest that the master biofilm regulator, CsgD, can modulate the virulence capacity of *S.* Typhimurium.

Bet hedging is characterized as a risk-spreading strategy displayed by clonal populations, where each phenotype performs more or less well at any time, depending on the selection pressures present (80). Although we can't be certain of the evolutionary selections that led to bistable CsgD expression, the passage of *Salmonella* isolates between host and non-host environments represents a model of unpredictability (7). Aggregation is known to have many evolutionary benefits to ensure the survival of populations, but instead of it being interpreted as altruistic behavior, it may be driven by harsh penalties associated with an individual not being part of the group (81). For pathogens like *S.* Typhimurium, the maintenance of a subset of cells primed for invasion but not ideally suited for environmental survival would be a penalty unless a host is encountered. The energy commitment required to make SPI-1 T3SS organelles (82) or extracellular matrix polymers (83, 84) in different subpopulations indicates that it might be a strategy for *S.* Typhimurium to prepare for unpredictability. A proportion of cells could survive exposure to either the host or non-host environments, with the ultimate goal being preservation of the shared genome for the next generation (54, 80). We speculate that signaling can occur between the host immune system, the microbiota and the invading pathogen, with the distinct possibility that *S.* Typhimurium, and presumably other NTS isolates, can modulate their virulence (85) or persistence (7) as a result of these interactions.

As pathogens evolve, there is thought to be continual selection pressure for them to adapt to their hosts (86). Host-adaptation has the benefit of reducing the variation and unpredictability that a bacterial pathogen is exposed to, but makes strains more dependent

on a direct means of transmission (4). Differences in transmission could potentially explain the correlation between an invasive, host-adapted way of life and loss of the rdar morphotype in *Salmonella* and *E. coli* (17, 26). In *E. coli*, bistable expression of CsgD does occur in specific pathogenic strains (87); however, human-adapted enteroinvasive *E. coli* and *Shigella* isolates have lost their ability to aggregate (27). Similarly, the evolution of *S.* Typhi has been strongly influenced by the human carrier state (5) and nearly all isolates are negative for the rdar morphotype (25, 26). We hypothesize that host-adapted or host-restricted isolates favor a single cell mode of growth because shedding from chronic carriers makes them less dependent on long-term survival to complete the transmission cycle. *S.* Typhi is capable of other forms of biofilm growth that are required for persistence inside human carriers, such as the attachment to the surfaces of gallstones (88) and recent evidence indicates that *S.* Typhi can be transmitted indirectly via contaminated water (89). This suggests that there are other factors to consider when correlating a loss of persistence with host-adaptation. Nevertheless, identifying the molecular mechanisms that connect virulence and persistence may hold the key for understanding transmission and ultimately reducing the spread of these human pathogens. We predict that the bet-hedging program described here for *S.* Typhimurium is a common strategy for survival, especially for other enteric pathogens that spend a significant portion of their lifecycle in the environment.

4.7 Acknowledgements

This research was supported by Natural Sciences and Engineering Research Council (NSERC) Discovery grants to APW (#386063-2010) and ADSC (#435784-2013) and through the Jarislowsky Chair in Biotechnology to APW. KDM was supported by a Canada Graduate Scholarship from NSERC, YW by a Postdoctoral Research Award from the Saskatchewan Health Research Foundation, and CSW by an Undergraduate Research Award from the University of Saskatchewan. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

The authors are grateful to Don Wilson, Stew Walker and the VIDO animal care staff for professional help with the animal experiments; to Diane Miller, Yongjun Zhao and the research team at Canada's Michael Smith Genome Sciences Centre for RNA-seq;

to Jun Han and Derek Smith at the UVic Genome BC Proteomics Centre for c-di-GMP measurements; to Alex Pasternak for assistance with confocal microscopy; to Anjuman Ara, Neil Rawlyk, and Wayne Connor for laboratory assistance; and to Mike Surette, Scott Napper, Peter Howard and Jos Vanderleyden for helpful discussions.

4.8 References

1. **Majowicz SE, Musto J, Scallan E, Angulo FJ, Kirk M, O'Brien SJ, Jones TF, Fazil A, Hoekstra RM.** 2010. The global burden of nontyphoidal *Salmonella* gastroenteritis. Clin Infect Dis 50: 882-889.

2. **Winter SE, Thiennimitr P, Winter MG, Butler BP, Huseby DL, Crawford RW, Russell JM, Bevins CL, Adams LG, Tsolis RM, Roth JR, Baumler AJ.** 2010. Gut inflammation provides a respiratory electron acceptor for *Salmonella*. Nature **467:**426-429.

3. **Winfield MD, Groisman EA.** 2003. Role of nonhost environments in the lifestyles of *Salmonella* and *Escherichia coli*. Appl Environ Microbiol **69:**3687-3694.

4. **Blaser MJ, Kirschner D.** 2007. The equilibria that allow bacterial persistence in human hosts. Nature **449:**843-849.

5. **Holt KE, Parkhill J, Mazzoni CJ, Roumagnac P, Weill F-X, Goodhead I, Rance R, Baker S, Maskell DJ, Wain J, Dolecek C, Achtman M, Dougan G.** 2008. High-throughput sequencing provides insights into genome variation and evolution in *Salmonella* Typhi. Nature Genet **40:**987-993.

6. **Moest TP, Meresse S.** 2013. *Salmonella* T3SSs: successful mission of the secret(ion) agents. Curr Opin Microbiol **16:**38-44.

7. **Monack DM.** 2012. *Salmonella* persistence and transmission strategies. Curr Opin Microbiol **15:**100-107.

8. **Waldner LL, MacKenzie KD, Koster W, White AP.** 2012. From Exit to Entry: Long-term Survival and Transmission of *Salmonella*. Pathogens **1:**128-155.

9. **Ziehm D, Dreesman J, Campe A, Kreienbrock L, Pulz M.** 2013. Risk factors associated with sporadic salmonellosis in adults: a case-control study. Epidemiol Infect **141:**284-292.

10. **Collinson SK, Emody L, Muller KH, Trust TJ, Kay WW.** 1991. Purification and characterization of thin, aggregative fimbriae from *Salmonella enteritidis*. J Bacteriol **173:**4773-4781.

11. **Romling U, Sierralta WD, Eriksson K, Normark S.** 1998. Multicellular and aggregative behaviour of *Salmonella typhimurium* strains is controlled by mutations in the *agfD* promoter. Mol Microbiol **28:**249-264.

12. **Gibson DL, White AP, Snyder SD, Martin S, Heiss C, Azadi P, Surette M, Kay WW.** 2006. *Salmonella* produces an O-antigen capsule regulated by AgfD and important for environmental persistence. J Bacteriol **188:**7722-7730.

13. **Zogaj X, Nimtz M, Rohde M, Bokranz W, Romling U.** 2001. The multicellular morphotypes of *Salmonella typhimurium* and *Escherichia coli* produce cellulose as the second component of the extracellular matrix. Mol Microbiol **39:**1452-1463.

14. **Doran JL, Collinson SK, Burian J, Sarlos G, Todd EC, Munro CK, Kay CM, Banser PA, Peterkin PI, Kay WW.** 1993. DNA-based diagnostic tests for *Salmonella* species targeting *agfA*, the structural gene for thin, aggregative fimbriae. J Clin Microbiol **31:**2263-2273.

15. **Solano C, Garcia B, Valle J, Berasain C, Ghigo JM, Gamazo C, Lasa I.** 2002. Genetic analysis of *Salmonella enteritidis* biofilm formation: critical role of cellulose. Mol Microbiol **43:**793-808.

16. **Romling U, Bian Z, Hammar M, Sierralta WD, Normark S.** 1998. Curli fibers are highly conserved between *Salmonella typhimurium* and *Escherichia coli* with respect to operon structure and regulation. J Bacteriol **180:**722-731.

17. **White AP, Sibley KA, Sibley CD, Wasmuth JD, Schaefer R, Surette MG, Edge TA, Neumann NF.** 2011. Intergenic sequence comparison of *Escherichia coli* isolates reveals lifestyle adaptations but not host specificity. Appl Environ Microbiol **77:**7620-7632.

18. **Anriany YA, Weiner RM, Johnson JA, De Rezende CE, Joseph SW.** 2001. *Salmonella enterica* serovar Typhimurium DT104 displays a rugose phenotype. Appl Environ Microbiol **67:**4048-4056.

19. **Scher K, Romling U, Yaron S.** 2005. Effect of heat, acidification, and chlorination on *Salmonella enterica* serovar Typhimurium cells in a biofilm formed at the air-liquid interface. Appl Environ Microbiol **71:**1163-1168.

20. **Uhlich GA, Cooke PH, Solomon EB.** 2006. Analyses of the red-dry-rough phenotype of an *Escherichia coli* O157:H7 strain and its role in biofilm formation and resistance to antibacterial agents. Appl Environ Microbiol **72:**2564-2572.

21. **White AP, Gibson DL, Kim W, Kay WW, Surette MG.** 2006. Thin aggregative fimbriae and cellulose enhance long-term survival and persistence of *Salmonella*. J Bacteriol **188:**3219-3227.

22. **Austin JW, Sanders G, Kay WW, Collinson SK.** 1998. Thin aggregative fimbriae enhance *Salmonella enteritidis* biofilm formation. FEMS Microbiol Lett **162:**295-301.

23. **Barak JD, Gorski L, Naraghi-Arani P, Charkowski AO.** 2005. *Salmonella enterica* virulence genes are required for bacterial attachment to plant tissue. Appl Environ Microbiol **71:**5685-5691.

24. **Apel D, White AP, Grassl GA, Finlay BB, Surette MG.** 2009. Long-term survival of *Salmonella enterica* serovar Typhimurium reveals an infectious state that is underrepresented on laboratory media containing bile salts. Appl Environ Microbiol **75:**4923-4925.

25. **White AP, Gibson DL, Grassl GA, Kay WW, Finlay BB, Vallance BA, Surette MG.** 2008. Aggregation via the red, dry, and rough morphotype is not a virulence adaptation in *Salmonella enterica* serovar Typhimurium. Infect Immun **76:**1048-1058.

26. **Romling U, Bokranz W, Rabsch W, Zogaj X, Nimtz M, Tschape H.** 2003. Occurrence and regulation of the multicellular morphotype in *Salmonella* serovars important in human disease. Intl J Med Microbiol **293:**273-285.

27. **Sakellaris H, Hannink NK, Rajakumar K, Bulach D, Hunt M, Sasakawa C, Adler B.** 2000. Curli loci of *Shigella* spp. Infect Immun **68:**3780-3783.

28. **Ahmad I, Wigren E, Le Guyon S, Vekkeli S, Blanka A, El Mouali Y, Anwar N, Chuah ML, Lunsdorf H, Frank R, Rhen M, Liang ZX, Lindqvist Y, Romling U.** 2013. The EAL-like protein STM1697 regulates virulence phenotypes, motility and biofilm formation in *Salmonella* Typhimurium. Mol Microbiol **90:**1216-1232.

29. **Hufnagel DA, Tukel C, Chapman MR.** 2013. Disease to dirt: the biology of microbial amyloids. PLoS Pathog **9:**e1003740.

30. **Grantcharova N, Peters V, Monteiro C, Zakikhany K, Romling U.** 2010. Bistable expression of CsgD in biofilm development of *Salmonella enterica* serovar Typhimurium. J Bacteriol **192:**456-466.

31. **Steenackers H, Hermans K, Vanderleyden J, De Keersmaecker SCJ.** 2012. *Salmonella* biofilms: An overview on occurrence, structure, regulation and eradication. Food Res Intl **45:**502-531.

32. **Datsenko KA, Wanner BL.** 2000. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. Proc Natl Acad Sci USA **97:**6640-6645.

33. **Desin TS, Lam PK, Koch B, Mickael C, Berberov E, Wisner AL, Townsend HG, Potter AA, Koster W.** 2009. *Salmonella enterica* serovar Enteritidis

pathogenicity island 1 is not essential for but facilitates rapid systemic spread in chickens. Infect Immun **77:**2866-2875.

34. **Maloy SR, Stewart VJ, Taylor RK.** 1996. Genetic analysis of pathogenic bacteria: a laboratory manual. Cold Spring Harbor Laboratory Press, Plainview, NY.

35. **White AP, Weljie AM, Apel D, Zhang P, Shaykhutdinov R, Vogel HJ, Surette MG.** 2010. A global metabolic shift is linked to *Salmonella* multicellular development. PLoS ONE **5:**e11814.

36. **Choi KH, Gaynor JB, White KG, Lopez C, Bosio CM, Karkhoff-Schweizer RR, Schweizer HP.** 2005. A Tn7-based broad-range bacterial cloning and expression system. Nat Methods **2:**443-448.

37. **Bobrov AG, Kirillina O, Ryjenkov DA, Waters CM, Price PA, Fetherston JD, Mack D, Goldman WE, Gomelsky M, Perry RD.** 2011. Systematic analysis of cyclic di-GMP signalling enzymes and their role in biofilm formation and virulence in *Yersinia pestis*. Mol Microbiol **79:**533-551.

38. **Parkhomchuk D, Borodina T, Amstislavskiy V, Banaru M, Hallen L, Krobitsch S, Lehrach H, Soldatov A.** 2009. Transcriptome analysis by strand-specific sequencing of complementary DNA. Nucleic Acids Res **37:**e123.

39. **Jarvik T, Smillie C, Groisman EA, Ochman H.** 2010. Short-term signatures of evolutionary change in the *Salmonella enterica* serovar Typhimurium 14028 genome. J Bacteriol **192:**560-567.

40. **Kroger C, Dillon SC, Cameron AD, Papenfort K, Sivasankaran SK, Hokamp K, Chao Y, Sittka A, Hebrard M, Handler K, Colgan A, Leekitcharoenphon P, Langridge GC, Lohan AJ, Loftus B, Lucchini S, Ussery DW, Dorman CJ, Thomson NR, Vogel J, Hinton JC.** 2012. The transcriptional landscape and small RNAs of *Salmonella enterica* serovar Typhimurium. Proc Natl Acad Sci U S A **109:**E1277-1286.

41. **Robinson MD, McCarthy DJ, Smyth GK.** 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics **26:**139-140.

42. **Benjamini Y, Hochberg Y.** 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J Roy Stat Soc B **57:**289-300.

43. **Robinson MD, Smyth GK.** 2008. Small-sample estimation of negative binomial dispersion, with applications to SAGE data. Biostatistics **9:**321-332.

44. **Romling U, Galperin MY, Gomelsky M.** 2013. Cyclic di-GMP: the first 25 years of a universal bacterial second messenger. Microbiol Mol Biol Rev **77:**1-52.

45. **Garcia B, Latasa C, Solano C, Garcia-del Portillo F, Gamazo C, Lasa I.** 2004. Role of the GGDEF protein family in *Salmonella* cellulose biosynthesis and biofilm formation. Mol Microbiol **54:**264-277.

46. **Simm R, Morr M, Kader A, Nimtz M, Romling U.** 2004. GGDEF and EAL domains inversely regulate cyclic di-GMP levels and transition from sessility to motility. Mol Microbiol **53:**1123-1134.

47. **Le Guyon S, Simm R, Rhen M, Romling U.** 2014. Dissecting the c-di-GMP signaling network regulating motility in *Salmonella enterica* serovar Typhimurium. Environ Microbiol doi:10.1111/1462-2920.12580.

48. **Ahmad I, Lamprokostopoulou A, Le Guyon S, Streck E, Barthel M, Peters V, Hardt WD, Romling U.** 2011. Complex c-di-GMP signaling networks mediate transition between virulence properties and biofilm formation in *Salmonella enterica* serovar Typhimurium. PLoS ONE **6:**e28351.

49. **Solano C, Garcia B, Latasa C, Toledo-Arana A, Zorraquino V, Valle J, Casals J, Pedroso E, Lasa I.** 2009. Genetic reductionist approach for dissecting individual roles of GGDEF proteins within the c-di-GMP signaling network in *Salmonella*. Proc Natl Acad Sci USA **106:**7997-8002.

50. **Rouf SF, Ahmad I, Anwar N, Vodnala SK, Kader A, Romling U, Rhen M.** 2011. Opposing contributions of polynucleotide phosphorylase and the membrane protein NlpI to biofilm formation by *Salmonella enterica* serovar Typhimurium. J Bacteriol **193:**580-582.

51. **Ryjenkov DA, Simm R, Romling U, Gomelsky M.** 2006. The PilZ domain is a receptor for the second messenger c-di-GMP: the PilZ domain protein YcgR controls motility in enterobacteria. J Biol Chem **281:**30310-30314.

52. **Zorraquino V, Garcia B, Latasa C, Echeverz M, Toledo-Arana A, Valle J, Lasa I, Solano C.** 2013. Coordinated cyclic-di-GMP repression of *Salmonella* motility through YcgR and cellulose. J Bacteriol **195:**417-428.

53. **Simm R, Remminghorst U, Ahmad I, Zakikhany K, Romling U.** 2009. A role for the EAL-like protein STM1344 in regulation of CsgD expression and motility in *Salmonella enterica* serovar Typhimurium. J Bacteriol **191:**3928-3937.

54. **Dubnau D, Losick R.** 2006. Bistability in bacteria. Mol Microbiol **61:**564-572.

55. **Siebring J, Sorg RA, Herber M, Kuipers OP.** 2012. Take it or Leave it: Mechanisms Underlying Bacterial Bistable Regulatory Networks, p 305-332. *In* Filloux AAM (ed), Bacterial Regulatory Networks. Caister Academic Press, Norfolk, UK.

56. **Gualdi L, Tagliabue L, Landini P.** 2007. Biofilm formation-gene expression relay system in *Escherichia coli*: modulation of sigmaS-dependent gene

expression by the CsgD regulatory protein via sigmaS protein stabilization. J Bacteriol **189:**8034-8043.

57. **Golubeva YA, Sadik AY, Ellermeier JR, Slauch JM.** 2012. Integrating global regulatory input into the *Salmonella* pathogenicity island 1 type III secretion system. Genetics **190:**79-90.

58. **Ono S, Goldberg MD, Olsson T, Esposito D, Hinton JC, Ladbury JE.** 2005. H-NS is a part of a thermally controlled mechanism for bacterial gene regulation. Biochem J **391:**203-213.

59. **Galan JE, Curtiss R, 3rd.** 1990. Expression of *Salmonella typhimurium* genes required for invasion is regulated by changes in DNA supercoiling. Infect Immun **58:**1879-1885.

60. **Kroger C, Colgan A, Srikumar S, Handler K, Sivasankaran SK, Hammarlof DL, Canals R, Grissom JE, Conway T, Hokamp K, Hinton JC.** 2013. An infection-relevant transcriptomic compendium for *Salmonella enterica* serovar Typhimurium. Cell Host Microbe **14:**683-695.

61. **Sanowar S, Singh P, Pfuetzner RA, Andre I, Zheng H, Spreter T, Strynadka NC, Gonen T, Baker D, Goodlett DR, Miller SI.** 2010. Interactions of the transmembrane polymeric rings of the *Salmonella enterica* serovar Typhimurium type III secretion system. mBio **1:**e00158-00110.

62. **Lara-Tejero M, Galan JE.** 2009. *Salmonella enterica* serovar Typhimurium pathogenicity island 1-encoded type III secretion system translocases mediate intimate attachment to nonphagocytic cells. Infect Immun **77:**2635-2642.

63. **Mirold S, Ehrbar K, Weissmuller A, Prager R, Tschape H, Russmann H, Hardt WD.** 2001. *Salmonella* host cell invasion emerged by acquisition of a mosaic of separate genetic elements, including *Salmonella* pathogenicity island 1 (SPI1), SPI5, and *sopE2*. J Bacteriol **183:**2348-2358.

64. **Misselwitz B, Barrett N, Kreibich S, Vonaesch P, Andritschke D, Rout S, Weidner K, Sormaz M, Songhet P, Horvath P, Chabria M, Vogel V, Spori DM, Jenny P, Hardt WD.** 2012. Near surface swimming of *Salmonella* Typhimurium explains target-site selection and cooperative invasion. PLoS Pathog **8:**e1002810.

65. **Mouslim C, Hughes KT.** 2014. The effect of cell growth phase on the regulatory cross-talk between flagellar and Spi1 virulence gene expression. PLoS Pathog **10:**e1003987.

66. **Osborne SE, Coombes BK.** 2011. Transcriptional priming of *Salmonella* Pathogenicity Island-2 precedes cellular invasion. PLoS ONE **6:**e21648.

67. **Brown NF, Vallance BA, Coombes BK, Valdez Y, Coburn BA, Finlay BB.** 2005. *Salmonella* pathogenicity island 2 is expressed prior to penetrating the intestine. PLoS Pathog **1:**e32.

68. **Galan JE, Curtiss R, 3rd.** 1989. Cloning and molecular characterization of genes whose products allow *Salmonella typhimurium* to penetrate tissue culture cells. Proc Natl Acad Sci USA **86:**6383-6387.

69. **Santiviago CA, Reynolds MM, Porwollik S, Choi SH, Long F, Andrews-Polymenis HL, McClelland M.** 2009. Analysis of pools of targeted *Salmonella* deletion mutants identifies novel genes affecting fitness during competitive infection in mice. PLoS Pathog **5:**e1000477.

70. **Tukel C, Nishimori JH, Wilson RP, Winter MG, Keestra AM, van Putten JP, Baumler AJ.** 2010. Toll-like receptors 1 and 2 cooperatively mediate immune responses to curli, a common amyloid from enterobacterial biofilms. Cell Microbiol **12:**1495-1505.

71. **Rapsinsky GJ, Wynosky-Dolfi MA, Oppong GO, Tursi SA, Wilson RP, Brodsky IE, Tukel C.** 2015. Toll-like receptor 2 and NLRP3 cooperate to recognize a functional bacterial amyloid, curli. Infect Immun **83:**693-701.

72. **Novick A, Weiner M.** 1957. Enzyme Induction as an All-or-None Phenomenon. Proc Natl Acad Sci USA **43:**553-566.

73. **Boehm A, Vogel J.** 2012. The *csgD* mRNA as a hub for signal integration via multiple small RNAs. Mol Microbiol **84:**1-5.

74. **Claessen D, Rozen DE, Kuipers OP, Sogaard-Andersen L, van Wezel GP.** 2014. Bacterial solutions to multicellularity: a tale of biofilms, filaments and fruiting bodies. Nat Rev Microbiol **12:**115-124.

75. **Ogasawara H, Yamamoto K, Ishihama A.** 2011. Role of the biofilm master regulator CsgD in cross-regulation between biofilm formation and flagellar synthesis. J Bacteriol **193:**2587-2597.

76. **Nielsen AT, Dolganov NA, Rasmussen T, Otto G, Miller MC, Felt SA, Torreilles S, Schoolnik GK.** 2010. A bistable switch and anatomical site control *Vibrio cholerae* virulence gene expression in the intestine. PLoS Pathog **6:**e1001102.

77. **Ackermann M, Stecher B, Freed NE, Songhet P, Hardt WD, Doebeli M.** 2008. Self-destructive cooperation mediated by phenotypic noise. Nature **454:**987-990.

78. **Diard M, Garcia V, Maier L, Remus-Emsermann MN, Regoes RR, Ackermann M, Hardt WD.** 2013. Stabilization of cooperative virulence by the expression of an avirulent phenotype. Nature **494:**353-356.

79. **Sturm A, Heinemann M, Arnoldini M, Benecke A, Ackermann M, Benz M, Dormann J, Hardt WD.** 2011. The cost of virulence: retarded growth of *Salmonella* Typhimurium cells expressing type III secretion system 1. PLoS Pathog **7:**e1002143.

80. **de Jong IG, Haccou P, Kuipers OP.** 2011. Bet hedging or not? A guide to proper classification of microbial survival strategies. Bioessays **33:**215-223.

81. **Parrish JK, Edelstein-Keshet L.** 1999. Complexity, pattern, and evolutionary trade-offs in animal aggregation. Science **284:**99-101.

82. **Ali SS, Soo J, Rao C, Leung AS, Ngai DH, Ensminger AW, Navarre WW.** 2014. Silencing by H-NS Potentiated the Evolution of *Salmonella*. PLoS Pathog **10:**e1004500.

83. **Davidson CJ, White AP, Surette MG.** 2008. Evolutionary loss of the rdar morphotype in *Salmonella* as a result of high mutation rates during laboratory passage. ISME J **2:**293-307.

84. **Smith DR, Chapman MR.** 2010. Economical evolution: microbes reduce the synthetic cost of extracellular proteins. mBio **1:**e00131-00110.

85. **Mastroeni P, Morgan FJ, McKinley TJ, Shawcroft E, Clare S, Maskell DJ, Grant AJ.** 2011. Enhanced virulence of *Salmonella enterica* serovar Typhimurium after passage through mice. Infect Immun **79:**636-643.

86. **Lenski RE, May RM.** 1994. The evolution of virulence in parasites and pathogens: reconciliation between two competing hypotheses. J Theor Biol **169:**253-265.

87. **DePas WH, Hufnagel DA, Lee JS, Blanco LP, Bernstein HC, Fisher ST, James GA, Stewart PS, Chapman MR.** 2013. Iron induces bimodal population development by *Escherichia coli*. Proc Natl Acad Sci USA **110:**2629-2634.

88. **Crawford RW, Rosales-Reyes R, Ramirez-Aguilar MdlL, Chapa-Azuela O, Alpuche-Aranda C, Gunn JS.** 2010. Gallstones play a significant role in *Salmonella* spp. gallbladder colonization and carriage. Proc Natl Acad Sci USA **107:**4353-4358.

89. **Baker S, Holt KE, Clements AC, Karkey A, Arjyal A, Boni MF, Dongol S, Hammond N, Koirala S, Duy PT, Nga TV, Campbell JI, Dolecek C, Basnyat B, Dougan G, Farrar JJ.** 2011. Combined high-resolution genotyping and geospatial analysis reveals modes of endemic urban typhoid fever transmission. Open Biol **1:**110008.

90. **Kosarewicz A, Konigsmaier L, Marlovits TC.** 2012. The blueprint of the type-3 injectisome. Philos Trans R Soc Lond B Biol Sci **367:**1140-1154.

5.0 INVASIVE NONTYPHOIDAL *SALMONELLA* STRAINS TYPHIMURIUM AND ENTERITIDIS ARE IMPAIRED FOR THE BIOFILM PHENOTYPE

Keith D. MacKenzie[1,2], Nancy L. Herman[2], Dakoda J. Herman[3], Yejun Wang[4], Nicholas A. Feasey[5,6], and Aaron P. White[1,2]*

[1] Vaccine and Infectious Disease Organization-International Vaccine Centre, Saskatoon, SK., Canada

[2] Department of Microbiology and Immunology, University of Saskatchewan, Saskatoon, SK., Canada

[3] Department of Biochemistry, University of Saskatchewan, Saskatoon, SK., Canada

[4] Department of Medical Genetics & Bioinformatics, Shenzhen University Health Science Center, Guangdong, China

[5] Department of Clinical Sciences, Liverpool School of Tropical Medicine, United Kingdom

[6] Malawi Liverpool Wellcome Trust Clinical Research Programme, University of Malawi College of Medicine, Blantyre, Malawi

* Corresponding author, e-mail: aaron.white@usask.ca

To be submitted to *Journal of Bacteriology*.

Author contributions:

Keith MacKenzie was responsible for experimental design, data generation, statistical analyses, and for writing the manuscript.

Nancy Herman and **Dakoda Herman** generated some of the plasmids and bacterial strains used in this study.

Yejun Wang provided the bioinformatics analysis that identified the nucleotide polymorphisms in *S.* Typhimurium D23580. He is also credited with the computational construction of the *S.* Enteritidis strains 4931 and 301 genome sequences.

Nicholas Feasey provided guidance with experimental design and edited the manuscript.

Aaron White assisted with experimental design, wrote the manuscript, and held the research grant that supported this work.

5.1 Interface

In sub-Saharan Africa, nontyphoidal *Salmonella* have surpassed *Salmonella* Typhi as the predominant cause of bacterial bloodstream infections. Little is known about the reservoirs and routes of transmission for nontyphoidal *Salmonella* strains associated with an invasive disease. In chapter 4, our characterization of multicellular aggregates and planktonic cells suggested that the subpopulations that result from CsgD bistability are part of a conserved transmission strategy used by *Salmonella*. In this chapter, we applied the principles we had learned about *Salmonella* biofilm formation to better understand the lifecycle of invasive nontyphoidal *Salmonella* strains. We evaluated select strains of invasive nontyphoidal *Salmonella* for CsgD bistability and the ability to form multicellular aggregates and planktonic cells. We provide evidence that invasive nontyphoidal *Salmonella* are impaired for biofilm formation and link this observation to strain-specific genetic polymorphisms.

5.2 Abstract

Nontyphoidal *Salmonella* strains are predominantly associated with gastroenteritis and diarrheal disease in humans, but recent reports have linked some clades to bloodstream infections in sub-Saharan Africa. It is hypothesized that this modern epidemic of invasive nontyphoidal *Salmonella* infections has stemmed from the geographical availability of an immune-compromised human population. *Salmonella* strains associated with extraintestinal diseases in humans are often host-adapted, infecting a narrowed range of susceptible host species. At the genetic level, these strains consistently exhibit features of genome degradation, where loss of gene function is associated with pathogen replication in the intestine and transmission via the fecal-oral route. Biofilm formation is suspected to have a role in *Salmonella* transmission, allowing cells to persist within the environment between host infections. Following the established correlation between host adaptation and loss of biofilm phenotype, we investigated the biofilm-forming ability of two invasive nontyphoidal *Salmonella* strains isolated from Malawi, *S*. Typhimurium D23580 and *S*. Enteritidis D7795. We assessed the phenotypic and genotypic conservation of the CsgD-regulated biofilm phenotype in these strains

compared to a panel of typical nontyphoidal *Salmonella* strains associated with gastroenteritis and host-restricted *S.* Typhi strains belonging to the H1 and H58 haplotypes. Both *S.* Typhimurium D23580 and *S.* Enteritidis D7795 demonstrated impairments in biofilm formation, which we attributed to specific genetic polymorphisms that are unique to each invasive nontyphoidal strain. We predict that biofilm impairment may reflect changes in these strains as they become more host-adapted and have a greater reliance on person-to-person transmission.

5.3 Introduction

Salmonella enterica are responsible for two contrasting diseases that inflict a significant burden on global health. More than half of the approximately 2600 *Salmonella* serovars identified to date belong to the subspecies *enterica*, which is associated with infections in warm-blooded hosts and is accountable for over 95% of human cases of salmonellosis (1, 2). Most of the serovars in subspecies *enterica* contribute to over 150 million annual cases of gastroenteritis, with serovars Typhimurium and Enteritidis responsible for the majority of these infections (3, 4). In immunocompetent individuals, nontyphoidal *Salmonella* infections involve the short-term colonization of the pathogen within the gastrointestinal tract, resulting in a localized inflammatory immune response accompanied by profuse diarrheal production. *Salmonella* serovars responsible for such cases of gastroenteritis are collectively referred to as nontyphoidal *Salmonella*. In contrast, a small number of serovars in subspecies *enterica* are known as typhoidal *Salmonella*, and are associated with an estimated 26.9 million annual cases of enteric fever (5). This distinct disease involves invasion of the extraintestinal compartment, often without the induction of inflammation or diarrhea (6). While such infections can occur asymptomatically, clinical manifestations of typhoidal *Salmonella* infections may include a persistent and gradual fever that elevates in a stepwise manner, as well as other symptoms such as headache, chills, nausea, coughing, malaise, or a rapid pulse (6, 7). Typhoidal *Salmonella* have the potential to persist asymptomatically within their human host over periods of several weeks to years as a result of the pathogen's intracellular association with monocytes and macrophages of the reticuloendothelial system and potential long-term colonization of the gall bladder (8).

In contradiction to this dichotomy, nontyphoidal *Salmonella* serovars have become increasingly responsible for bacterial bloodstream infections in sub-Saharan Africa (9). Like typhoidal *Salmonella* infections, invasive nontyphoidal *Salmonella* (iNTS) disease frequently lacks diarrheal symptoms, with febrile illness being the dominant clinical presentation in 95% of cases (10). This disease is responsible for an estimated 681,000 deaths per year, with nonspecific symptomology, multidrug resistance, and poor clinical outcomes despite correct diagnosis increasing the urgency to better understand this disease (9). A review of several clinical studies has revealed a significant association between invasive infections with nontyphoidal *Salmonella* and the presence of an immunocompromised population in the sub-Saharan region, particularly children with malnutrition or severe malaria and adults with advanced infections of human immunodeficiency virus (HIV) (9-11). Failure of the immune system to maintain the intestinal epithelial barrier or control intracellular *Salmonella* infections in these individuals provides a unique opportunity for nontyphoidal *Salmonella* serovars to exploit the availability of the systemic niche (10).

Nontyphoidal and typhoidal *Salmonella* serovars associated with extra-intestinal diseases are representative of the process of host adaptation, where events that affect a pathogen's evolution subsequently shift the dynamics of host-pathogen interactions (12). Most nontyphoidal *Salmonella* infect a broad range of host species and as such are considered host-generalist pathogens (13). While the exact chronology of events that resulted in host specialization for invasive nontyphoidal and typhoidal *Salmonella* infections remains elusive, the switch to a systemic niche and increased virulence during human infections has been correlated with significant genomic changes in these pathogens (12). *Salmonella* that survive and replicate in the systemic niche are faced with novel selection pressures that drive these genomic changes, including gene acquisition, genome degradation, and genetic rearrangements. These genetic changes are apparent in genome comparisons between strains associated with gastroenteritis versus invasive, systemic infections. Genome degradation, or the loss of gene function due to deleterious mutations or pseudogene formation, is a consistent genetic signature demonstrated in both typhoidal serovars *S.* Typhi and *S.* Paratyphi A and in clades of *S.* Typhimurium and *S.* Enteritidis associated with iNTS disease (14-19). Many of these gene mutations affect

metabolic processes involved in anaerobic respiration on unique carbon sources, a mechanism that is pivotal to the replication and outgrowth of *Salmonella* in the inflamed intestinal tract (20-22). Additional gene mutations affect adhesins and type three secretion systems, which are essential to successful intestinal colonization and replication. For nontyphoidal *Salmonella* associated with gastroenteritis, these factors are naturally selected due to their important role in enhancing pathogen transmission to the next host via the fecal-oral route (23).

In addition to colonization and replication, *Salmonella* must be able to successfully transmit to the next host to ensure its continued survival. For host-generalist strains of nontyphoidal *Salmonella*, maximal transmission potential is achieved through unique resource utilization in the inflamed intestinal niche and adoption of a broad range of susceptible host species. As a consequence of the fecal-oral route of transmission, nontyphoidal *Salmonella* are exposed to the environment between host colonization events (23, 24). Biofilm formation is proposed to aid in the survival and persistence of nontyphoidal *Salmonella* cells during this environmental phase of the transmission cycle (25, 26). *Salmonella* biofilm physiology has been termed the rdar (red, dry, and rough) morphotype, where a self-produced extracellular matrix of curli fimbriae, cellulose, and O-antigen capsule polymers interconnects cells and facilitates their adherence to abiotic and biotic surfaces (27-30). Conditions of low temperatures, low osmolarity, and nutrient limitation promote the expression of the rdar morphotype by activating RpoS, the sigma factor for the general stress response, and CsgD, a member of the RpoS regulon and the primary transcriptional activator of the rdar morphotype (28, 29). The genes for both curli fimbriae and cellulose are highly conserved in *Salmonella* and *Escherichia coli*; however, almost all typhoidal *Salmonella* serovars are phenotypically negative for the rdar morphotype (31-36). As host-adapted pathogens, typhoidal serovars have potentially lost this mechanism of environmental survival, instead relying on chronic persistence in the human host and intermittent shedding via the biliary tract and gall bladder to improve their odds of transmission (37). Therefore, loss of the rdar morphotype may represent an additional signature of host adaptation.

Humans are also speculated to be the primary reservoir for invasive nontyphoidal Salmonellae (17, 19, 38). Preliminary field studies into the transmission of iNTS serovars

have emphasized the importance of person-to-person spread. In these studies, iNTS isolates retrieved from index patients were matched to isolates from persons in close contact, but not to isolates obtained from nearby environmental reservoirs such as animals, soil, sewer, water, or food products (38, 39). However, the level of host restriction for iNTS serovars is currently under debate. *S.* Typhimurium D23580, a strain of invasive *Salmonella* belonging to the predominant sequence type ST313, has been demonstrated to cause infection in different animal challenge models, including mice, cattle, chicks, and rhesus macaques (18, 40-43). In contrast, *S.* Enteritidis D7795, a representative of the recently described epidemic Central/Eastern African clade for this serovar, demonstrated a reduced ability to cause systemic infection in chicks (19). To date, there have been few investigations in the biofilm phenotype of these invasive nontyphoidal *Salmonella* strains.

In this study, we present a detailed evaluation of the biofilm-forming capacity of the type-strains of iNTS, *S.* Typhimurium D23580 and *S.* Enteritidis D7795, in comparison with typical NTS strains as well as host-restricted *S.* Typhi strains of the H1 and H58 haplotypes. We analyzed the rdar morphotype via growth on agar plates and multicellular aggregation in liquid cultures, both of which have been demonstrated to correlate with the ability of a strain to form biofilm. We monitored expression of RpoS, CsgD, and other biofilm transcriptional regulators in each strain, and searched for sequence polymorphisms in the genome sequences. *S.* Typhimurium D23580 demonstrated an impaired ability to aggregate both by the rdar morphotype and in liquid cultures, and *S.* Enteritidis D7795 was negative in both growth formats. In agreement with this, *S.* Typhimurium D23580 did produce curli fimbriae and cellulose at a reduced capacity, while D7795 was negative for both polymers. We identified a pivotal mutation in the promoter region of *csgD* in *S.* Enteritidis D7795, which resulted in this being a null promoter when compared to other nontyphoidal *Salmonella* isolates. For *S.* Typhimurium D23580, the changes were subtler, with a potential polymorphism detected in the STM1987 gene that produces the key biofilm signaling molecule, cyclic-di-GMP. Overall, our results correspond with the hypothesis of reduced host ranges for these two iNTS isolates, and suggests that mutations affecting the rdar morphotype may serve as a

sensitive indicator of the position of *Salmonella* strains on the continuum of host adaptation.

5.4 Materials and Methods

5.4.1 Bacterial strains, media, and growth conditions

The bacterial strains used in this study are listed in Table 5.1. For standard growth, strains were inoculated from frozen stocks onto LB agar (lysogeny broth-Miller plus 1.5% agar) and grown overnight at 37˚C. One isolated colony was used to inoculate 5 mL LB broth (containing 1% NaCl) and the culture was incubated for 18 hours at 37˚C with agitation at 200 RPM. For colony morphology assays, overnight cultures of each strain were normalized to an optical density of 1.0 at 600 nm and 2 µL were spotted onto 1% tryptone medium containing 1.5% Difco agar (T agar) (27). To visualize cellulose production, T agar was supplemented with calcofluor white (fluorescent brightener 28; Sigma-Aldrich Canada) at a final concentration of 200 µg/mL (34). To analyze liquid culture growth under biofilm-inducing conditions, approximately 1 x 10^9 CFU were inoculated into 100 mL of 1% tryptone, pH 7.4, and incubated at 28˚C for 24 or 48 hours with agitation at 200 rpm.

For bioluminescence assays, overnight cultures of the *Salmonella* strains were diluted 1 in 600 in 1% tryptone broth supplemented with 50 µg mL^{-1} Kn in a final volume of 150 µL in 96-well clear-bottom black plates (9520 Costar; Corning Inc.). To minimize evaporation of the medium during the assay, cultures were overlaid with 50 µL of mineral oil. Cultures were assayed for absorbance (590 to 600 nm, 0.1 s) and luminescence (1s; in counts per second [CPS]) every 30 min during growth at 28˚C with agitation in a Victor X3 multilabel plate reader (Perkin-Elmer).

5.4.2 Generation of bacterial luciferase reporters and pBR322-STM1987 plasmid vectors

The pCS26 bacterial luciferase (*lux*) operon fusion plasmids containing *csgDEFG*, *csgBAC*, *adrA*, and *rpoS* promoter sequences from *S.* Typhimurium 14028 have been described previously (26, 32). The RpoS-dependent reporter plasmid sig38H4 contains the *luxCDABE* operon preceded by a promoter designed based on the alignment

Table 5.1. Strains used in this study.

Strain or Plasmid	Genotype	Source or Reference
Strains		
S. Typhimurium 14028	Wild-type strain	ATCC
S. Typhimurium SL1344	Wild-type strain	ATCC (Wolfgang Köster, University of Saskatchewan)
S. Enteritidis ATCC 4931	Wild-type strain	ATCC (Darren Korber, University of Saskatchewan)
S. Enteritidis 301 (Sal18)	Wild-type strain	Wolfgang Köster, University of Saskatchewan (88)
S. Typhimurium D23580	Wild-type strain	Gordon Dougan (17)
S. Enteritidis D7795	Wild-type strain	Gordon Dougan (19)
S. Typhi E02-2759	Wild-type strain	Gordon Dougan (16)
S. Typhi E03-9804	Wild-type strain	Gordon Dougan (16)
S. Typhi ISP03-07467	Wild-type strain	Gordon Dougan (16)
S. Typhi ISP04-06979	Wild-type strain	Gordon Dougan (16)
S. Typhi 8(04)N	Wild-type strain	Gordon Dougan (16)
S. Typhi CT18	Wild-type strain	Gordon Dougan
S. Typhimurium 14028	Δ*STM1987*	This study
Plasmids		
pCS26		(44)
pCS26-adrA		(26)
pCS26-csgB		(26)
pCS26-sig38H4		(26)
pU220		(44)
pU220-csgD		(26)
pCS26-cpxR		This study
pKD3		(47)
pKD46		(47)
pBR322		
pBR322-STM1987 (14028)		This study
pBR322-STM1987 (D23580)		This study

of several RpoS-controlled promoters (26) To generate the pCS26-cpxR promoter-*luxCDABE* construct, the *cpx* intergenic region was PCR amplified from *S.* Typhimurium 14028 genomic DNA using primers cpxR1 and cpxR2 (Table 5.2) and Phusion high-fidelity DNA polymerase (New England BioLabs), with reaction conditions outlined by the manufacturer. The desired PCR product was purified (FroggaBio Inc.), sequentially digested with *Xho*I and *Bam*HI (New England BioLabs), and ligated using T4 DNA ligase (New England BioLabs) into the pCS26 vector (XhoI-BamHI) containing the *luxCDABE* operon from *Photorhabdus luminescens* (44). To generate *csgDEFG* and *csgBAC* promoter-*lux* reporters, the *csg* intergenic region was PCR amplified from genomic DNA using primers agfD1 and agfD2 (Table 5.2). The PCR products were then ligated into either pCS26-Pac (XhoI-BamHI) or pU220 (BamHI-XhoI) to generate the *csgBAC* and *csgDEFG* promoter-reporter constructs, respectively. To generate pBR322-STM1987 plasmid vectors, DNA fragments containing *STM1987* with native promoters were first PCR amplified from *S.* Typhimurium 14028 or *S.* Typhimurium D23580 from genomic DNA using primers STM1987forEco and STM1987revAatII (Table 5.2). These PCR products were purified, sequentially digested with *Eco*RI and *Aat*II, and ligated using T4 DNA ligase into digested pBR322. Plasmids were transformed into *Salmonella* strains by electroporation and selected by growth at 37˚C on LB agar supplemented with either 50 μg mL^{-1} kanamycin (pCS26, pU220) or 10 μg mL^{-1} tetracycline (pBR322).

5.4.3 SDS-PAGE and Western blotting

Approximately 5 x 10^{10} CFU of planktonic cells were sedimented by centrifugation (11,000 x g; 10 min; 4˚C) and resuspended in 1 mL of SDS-PAGE sample buffer (without 2-mercaptoethanol and bromophenol blue). Approximately 30 mg samples of multicellular aggregates were resuspended in 1 mL of water and homogenized with a glass tissue grinder for 25 dounces. The cell material was sedimented (10,000 x g; 1 min) before resuspension in 1 mL of SDS-PAGE sample buffer. All samples were boiled for 5 min. Using the DC protein assay (Bio-Rad Laboratories), cell lysates were normalized to a final protein concentration of 3 mg/mL. Bromophenol blue (0.0002% final concentration) and 2-mercaptoethanol (0.2% final concentration) were added to each lysate before loading 15 μg of total protein per lane. SDS-PAGE was performed

Table 5.2. Oligonucleotides used in this study.

Primer	Sequence (5′ - 3′)[a]	Purpose
agfD1	GTGCTCGAGGGACTTCATTAA ACATGATG	To amplify the *csg* intergenic and untranslated regions from chromosomal DNA
agfD2	GCCGGATCCTGTTTTTCATGCT GTCAC	To amplify the *csg* intergenic and untranslated regions from chromosomal DNA
cpxR1	GCCCTCGAGGTAACTTTGCGC ATCGCTTG	To amplify *cpxRA* promoter region from chromosomal DNA
cpxR2	GCCGGATCCTTCATTGTTTACG TACCTCCG	To amplify *cpxRA* promoter region from chromosomal DNA
STM14_2408-ko_sense	GTGCCGCACGAAACACTGTTA GACAATCAAGGCTGGTTTAAA AAGCTGGCGTGTAGGCTGTAG CTGCTTC	To amplify *cat* gene product for lambda-red recombination
STM14_2408-ko_antisense	GAATGGACTATTTCTTTTCCCG CTCCTGAGTCGCGTCGCTGGC GCAAATACCTCCTTAGTTCCT ATTCCG	To amplify *cat* gene product for lambda-red recombination
STM1987forEco	GATCGAATTCAAACGGTGTTT CGCAC	To generate STM1987 promoter and gene insert from chromosomal DNA for pBR322 cloning
STM1987revAatII	GATCGACGTCGGACTATTTCT TTTCCCGCT	To generate STM1987 promoter and gene insert from chromosomal DNA for pBR322 cloning

[a] Nucleotide sequences corresponding to restriction enzyme sites are underlined.

with a 5% stacking gel and a 12 or 15% resolving gel. Proteins were transferred to nitrocellulose for 40 min at 25 V using a Trans-Blot SD semidry transfer cell (Bio-Rad Laboratories) in tris-glycine buffer supplemented with methanol. To detect curli fimbriae, cell debris was sedimented following boiling in SDS-PAGE sample buffer, then washed twice with 500 µL of distilled water, suspended in 250 µL of 90% formic acid, frozen at -80˚C, and lyophilized (27). The dried samples were resuspended in SDS-PAGE sample buffer and loaded directly without boiling into each SDS-PAGE gel lane (27). CsgD was detected using a CsgD-specific monoclonal antibody at a 1-in-6 dilution of tissue culture supernatant (ImmunoPrecise Antibodies Ltd., Victoria, BC) (45). To detect RpoS protein, a commercially available mouse polyclonal immune serum recognizing epitope 33 to 256 of *E. coli* sigma factor S protein was used at a 1-in-2000 dilution (BioLegend; 1RS1). CsgA, the major subunit of curli fimbriae, was detected by using a rabbit polyclonal serum raised against whole purified curli (27). GroEL was used as a protein-loading control and was detected with rabbit polyclonal immune serum at a 1-in-60,000 dilution (Sigma-Aldrich; G6532). Secondary antibodies IRDye® 800CW Goat anti-Mouse immunoglobulin G (IgG) or 680RD Goat anti-rabbit IgG were used at a 1-in-10,000 dilution and detected using the the Odyssey CLx imaging system and Image Studio 4.0 software package (Li-Cor Biosciences).

5.4.4 *Cis* versus *trans* reporter assays and statistical analysis

We evaluated strain-specific genetic changes to biofilm regulation using vector-based promoter-reporter fusion constructs. *Cis* assays allowed us to quantify the effect of strain-specific sequence changes within the promoter of a biofilm-related gene. In these assays, the reporter genes from *Photorhabdus luminescens* (*luxCDABE*) were fused to promoter sequences originating from each of the 12 strains included in this study. Expression of each promoter-reporter construct was monitored in the *S.* Typhimurium 14028 background. This experimental format ensures that promoter expression is measured in a consistent cellular environment, where the same collection of transcription factors is available, and their timing and expression are consistent. *Trans* assays measure the expression of a constant/specific promoter sequence within each strain to allow for detection of strain-specific variation in upstream regulatory components. In our

experiments, we transformed a collection of vectors with *S.* Typhimurium 14028 promoter sequences fused to *luxCDABE* into each strain included in this study. The expression of a given *S.* Typhimurium 14028 promoter sequence was then monitored in each strain background, thereby allowing for detection of strain-specific differences in the upstream regulatory environment affecting the expression of the given promoter.

Statistical analysis of the data was performed using GraphPad Prism version 7.0a. Data collected from *cis* and *trans* assays was reported as the maximum luciferase expression measured during the course of the experiment, and was expressed as the mean ± the standard deviation. This data was logarithmically transformed and evaluated for normal distribution using the Shapiro-Wilk normality test. If the data was normally distributed, comparisons of the mean maximum luciferase expression levels obtained from multiple experiments were performed using an ordinary one-way ANOVA with post-hoc analysis via Holm-Sidak's multiple comparisons test with statistical significance set at $p < 0.05$. If the data was not normally distributed, comparisons were performed using the Kruskal-Wallis test with post-hoc analysis via Dunn's multiple comparison test with statistical significant set at $p < 0.05$. For *cis* assays, the mean of maximal luciferase expression observed for a given strain was compared to the mean of every other strain in the experiment. For *trans* assays, the mean of maximal luciferase expression from a 14028 promoter-reporter construct in each strain was compared to expression of the construct in every other strain in the experiment.

5.4.5 Reference genome sequences

Whole genome sequences were obtained from the National Centre for Biotechnology Information (NCBI) via the following accession numbers: NC_016810 (*S.* Typhimurium SL1344); NC_016854 (*S.* Typhimurium D23580); NC_003198 (*S.* Typhi CT18). Mapped assemblies for *S.* Typhi H58 haplotype strains E02-2759, E03-9804, ISP03-07467, ISP04-06979 were available from http://www.sanger.ac.uk/Projects/S_typhi (16). Assembly of the *S.* Typhi 8(04)N was available from the European Nucleotide Archive under the name SGB112 and associated with the assembly number GCA_001362315.1. The sequence for *S.* Enteritidis D7795

was available from the Public Health England Pathogens BioProject on NCBI (accession number PRJNA248792).

5.4.6 Chromosomal DNA isolation

Overnight liquid cultures of *Salmonella* were sub-cultured 1:100 in 200 mL LB broth and grown at 37˚C for 2.5 hours. Approximately 7 x 10^8 cells were centrifuged (5465 x g, 10 minutes, 4˚C), resuspended in Tris-EDTA solution (25:1M, pH 8.0) to a total volume of 3.5 mL, and then treated with 10 mg lysozyme (Sigma-Aldrich; L6876) and 200 units of mutanolysin (Sigma-Aldrich; M9901) for 1 hour at 37˚C. For cell lysis, each sample was treated with 50 µL of 25% SDS, 1 mg proteinase K (Applied Biosystems; AM23548), and 125 µL of a 5M sodium chloride solution and incubated at 65˚C for 30 minutes. Nucleic acid was isolated from cell lysates using a series of phenol:chloroform:isoamyl and phenol:chloroform extractions, precipitated by the addition of ammonium acetate (at a final concentration of 2M), washed with ethanol, and resuspended in Tris-EDTA solution (10:1 M, pH 8.0). To remove RNA, RNase A was added to each sample at a final concentration of 0.2 mg/mL and incubated for 1 hour at 37˚C. Samples were purified once more by phenol:chloroform:isoamyl extraction, precipitated with ammonium acetate, washed with ethanol, and resuspended in a final volume of 200 µL of Tris-EDTA (10:1 M, pH 8.0).

5.4.7 Genome sequencing and sequence alignments

Purified chromosomal DNA samples from *S.* Enteritidis strains 4931 and 301 were fragmented by cup horn sonication with a high-intensity ultrasonic processor (Vibra-Cell, Danbury, CT) for 10 cycles of a 30-second pulses and 2 minute rest. DNA libraries were prepared from 1 µg of fragmented DNA using the NEBNext Ultra DNA Library Prep Kit for Illumina (New England BioLabs; E7645S) and NEBNext Multiplex Oligos for Illumina (Index Primer Set 2) (New England BioLabs; E7500S) according to the manufacturer's protocols. Adaptor-ligated DNA was size-selected between 400 and 500 bp total library size (length of insert sequence and adaptor sequence) following kit instructions. DNA samples were assessed for quality, purity, and integrity by using a NanoDrop ND-1000 spectrophotometer (Fisher Scientific) and an Agilent 2100

Bioanalyzer with a High Sensitivity DNA chip (Agilent Technologies; 5067-4626). To ensure correct adaptor ligation, samples were analyzed quantitative PCR using the KAPA Library Quantification Kit for the Illumina platform (KAPA Biosystems; KK4824). Samples were sequenced using the MiSeq Reagent Kit version 3, 600 cycles (2 x 300 bp read length) (Illumina; MS-102-3003). Sequencing reads were aligned to the *S.* Enteritidis reference genome P125109 using Geneious version 8.1.5.

Sequence alignments were performed in Geneious Pro version 8.1.5, using the ClustalW alignment and an IUB cost matrix (gap open cost of 15, gap extend cost of 6.66). A phylogenetic tree for strains included in this study was constructed based on the *csg* operon region by using the Geneious Tree Builder program included with the Geneious software package (46), with the Tamura-Nei model of genetic substitution and the neighbour-joining tree building method with bootstrapping (1000 replicates and support threshold of 70%).

5.4.8 Generation of *S.* Typhimurium 14028 Δ*STM1987* mutant strain

The *S.* Typhimurium 14028 Δ*STM1987* strain was generated using the lambda red recombinase knockout procedure (47). Primers containing a 50-nucleotide sequence on either side of the *STM1987* are listed in Table 5.2 and were used to amplify the *cat* gene from pKD3 using Phusion High-Fidelity DNA polymerase (New England BioLabs). The PCR products were purified (AxyPrep Mag PCR Cleanup kit; Axygen) and electroporated into *S.* Typhimurium 14028 cells containing pKD46. Mutant strains were first selected by growth at 37˚C on LB agar supplemented with 7 μg mL^{-1} chloramphenicol (Cm) before restreaking onto LB agar containing 30 μg mL^{-1} Cm. The mutations were moved into clean wild-type backgrounds by transduction with P22 phage (48). The *cat* gene was resolved from the chromosome using pCP20 (47). DNA sequencing of PCR products amplified from the chromosome of the Δ*STM1987* strain confirmed the loss of the majority of *STM1987*.

5.5 Results

5.5.1 Evaluation of the biofilm phenotype for invasive nontyphoidal and typhoidal *Salmonella* strains

The appearance of *Salmonella* macrocolonies grown on T agar has routinely been used as a diagnostic tool for the biofilm behaviour of a strain, where the solid, dry, and rough appearance of a biofilm-positive strain contrasts from the smooth, mucoidal texture of one with a biofilm-negative phenotype (27, 35). After 5 days of growth at 28˚C on T agar, nontyphoidal and typhoidal *Salmonella* strains were easily distinguishable based on the different appearances of their macrocolonies (Fig. 5.1). Control strains of nontyphoidal *Salmonella* associated with gastroenteritis in humans were positive for the rdar morphotype, as demonstrated by the formation of concentric rings and colony wrinkling on the macrocolony surface. Typhoidal *Salmonella* strains were rdar-negative in appearance, forming smooth, mucoid, and non-aggregative macrocolonies. *S.* Typhimurium D23580, an invasive nontyphoidal *Salmonella* strain, demonstrated a partial, incomplete biofilm pattern on its surface. Although macrocolonies of this strain appeared smooth and lacked the characteristic wrinkled appearance, there was evidence of textural differences between the outer edge of the colony, corresponding to the zone of active cell growth, and the inner region containing an older cell population (26, 49). These zones were visibly divided by a concentric ring pattern on the surface of the macrocolony. In contrast, *S.* Enteritidis D7795, the other representative invasive nontyphoidal strain included for study, was negative for the rdar morphotype.

Cellulose is an exopolysaccharide of the extracellular matrix that is essential for long-range interactions between cells in a rdar macrocolony (50). To detect the production of this polymer, macrocolonies of each strain were grown on T agar containing calcofluor white, a polysaccharide-binding fluorescent dye (34). Nontyphoidal *Salmonella* macrocolonies strongly produced cellulose, as demonstrated by their intense fluorescent signal under long-wave UV light. Similar fluorescence intensity was observed for *S.* Typhimurium D23580 macrocolonies, indicating the presence of cellulose production despite the partial rdar morphotype. In contrast, low fluorescent signaling

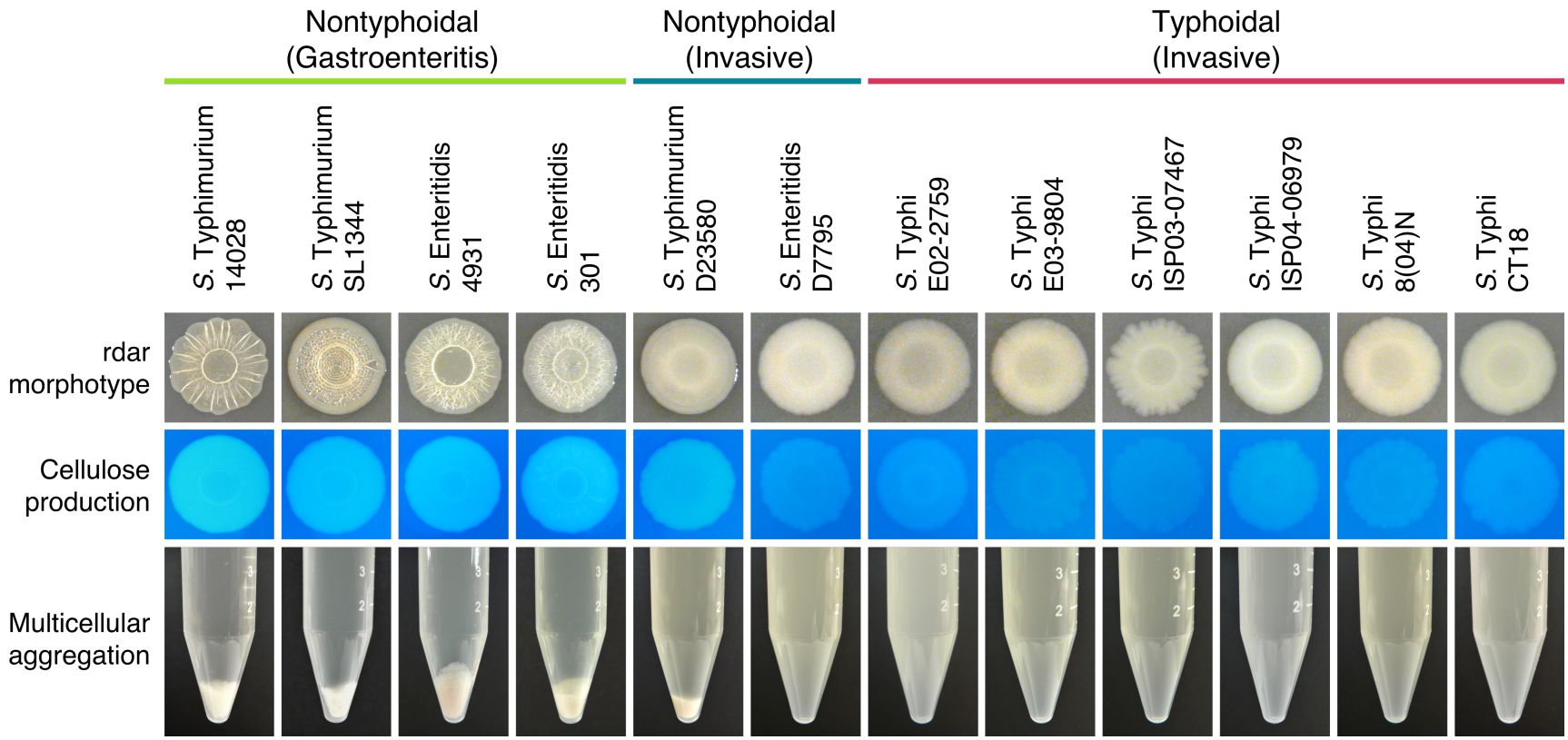

Figure 5.1. The biofilm phenotypes of nontyphoidal and typhoidal *Salmonella* strains.

Top panel: 2 µL of cells from overnight cultures were grown at 28˚C for 48 hours on T agar; the red, dry, and rough morphotype is indicative of biofilm formation, and presents as the formation of concentric rings and a wrinkled appearance on the surface of macrocolonies. Middle panel: Cells were grown for 48 hours at 28˚C on T agar supplemented with 200 µg mL^{-1} of calcofluor white dye. The white and fluorescent appearance of macrocolonies visualized under UV light is indicative of calcofluor binding to cellulose.

Bottom panel: Strains were grown in 1% tryptone broth for 48 hours at 28˚C; at the population level, bistable expression of *csgD* limits biofilm expression to a subpopulation of cells, resulting in phenotypic switching. This phenomenon is visualized by the presence of both multicellular aggregates and planktonic cells in aliquots from liquid cultures.

from both *S.* Enteritidis D7795 and typhoidal *Salmonella* macrocolonies indicated that cellulose production was comparably minimal or absent from these strains.

Macrocolony appearance and the rdar morphotype provide a population-level analysis of the biofilm phenotype. While strong biofilm-producing strains of *Salmonella* have an obvious biofilm phenotype, other strains that are moderate or partially impaired in biofilm production may be incorrectly categorized as biofilm-negative. It was recently demonstrated that biofilm-producing *Salmonella* strains grown in flask cultures will produce two unique subpopulations of multicellular aggregates and planktonic cells (45, 51). This heterogeneous expression of the biofilm phenotype is due to bistable expression of *csgD* within a clonal population. (52). We reasoned that multicellular aggregation in liquid cultures might provide a clearer diagnostic for the presence of a functioning biofilm phenotype in both strong and moderate biofilm-producing strains. Similar to the biofilm-positive nontyphoidal *Salmonella* strains included in our panel, cultures of *S.* Typhimurium D23580 contained visible multicellular aggregates and planktonic cell subpopulations. However, these aggregates were more slime-like in texture and made up a lower proportion of the population biomass compared to aggregates from nontyphoidal strains associated with gastroenteritis (data not shown). In contrast, cultures of *S.* Enteritidis D7795 and all typhoidal *Salmonella* strains were homogeneous and consisted solely of planktonic cells. Overall, our results suggested a strain-specific nature for the biofilm phenotype for invasive nontyphoidal *Salmonella* strains, where *S.* Typhimurium D23580 has retained a moderate but impaired biofilm phenotype, and *S.* Enteritidis has lost this phenotype altogether.

5.5.2 Expression of the *S.* Typhimurium 14028 biofilm gene promoter library in invasive nontyphoidal and typhoidal strains

Impairment or loss of the biofilm phenotype for invasive nontyphoidal *Salmonella* may reflect mutations in the transcriptional regulatory network that controls the expression of curli fimbriae and cellulose. To compare the functionality of the biofilm regulatory network in our panel of Salmonellae, we transformed each strain with a reporter library of biofilm-related gene promoters from *S.* Typhimurium 14028 fused to bacterial luciferase genes from *Photorhabdus luminescens*, *luxCDABE* (26, 32). For our

analysis of promoter activity, we compared the maximum expression of a given promoter in each strain relative to its expression in the native *S.* Typhimurium 14028 strain (Fig. 5.2). While our microplate culture conditions facilitated the growth of both nontyphoidal and typhoidal *Salmonella*, we noted a consistent delay in the exponential growth of *S.* Typhi strain E02-2759 (Fig. 5.2A).

We used transcription from the *cpxRA* promoter as an indicator of curli and cellulose biosynthesis (Figure 5.2B). Activation of the Cpx envelope stress response is proposed to represent a transition from the initial synthesis and establishment of biofilm polymers to a more mature biofilm state (53, 54). Envelope stress signals activate the two-component signaling system CpxAR, resulting in CpxR phosphorylation. In this activated state, CpxR potently inhibits further biosynthesis of curli fimbriae and cellulose by auto-activating its own transcription and repressing the transcription of both *csg* operons (53, 55-57). In our biofilm-positive nontyphoidal *Salmonella* strains, maximal expression of the *cpxRA* promoter ranged between 1 500 and 47,000 CPS. Expression of the *cpxRA* promoter in *S.* Typhimurium D23580 was consistent with its ability to form multicellular aggregates and further suggested the production of biofilm extracellular matrix components. In contrast, we observed a 2-log reduction in promoter activity for strains with a biofilm-negative phenotype.

CsgD promotes synthesis of the extracellular matrix through transcriptional activation of *csgBAC*, encoding the curli biosynthesis genes, and *adrA*, a diguanylate cyclase that indirectly activates cellulose biosynthesis through production of cyclic-di-GMP (29, 35, 58, 59). We observed a notable difference in the expression of either promoter between biofilm-positive and biofilm-negative *Salmonella* strains included in our study. White *et al.* previously reported *trans* variability in the expression of the *csgBAC* promoter between biofilm-positive strains of *Salmonella* (32). Consistent with this finding, the maximal expression of this promoter ranged between approximately 4,600 and 280,000 luminescence counts per second (CPS) in our biofilm-positive nontyphoidal *Salmonella* strains, with maximal expression of the promoter observed in its native strain, *S.* Typhimurium 14028 (Fig. 5.2C). Expression of the *csgBAC* promoter in invasive strain *S.* Typhimurium D23580 fell within this dynamic range. A similar pattern was observed for the *adrA* promoter; maximal expression ranged between 800 and

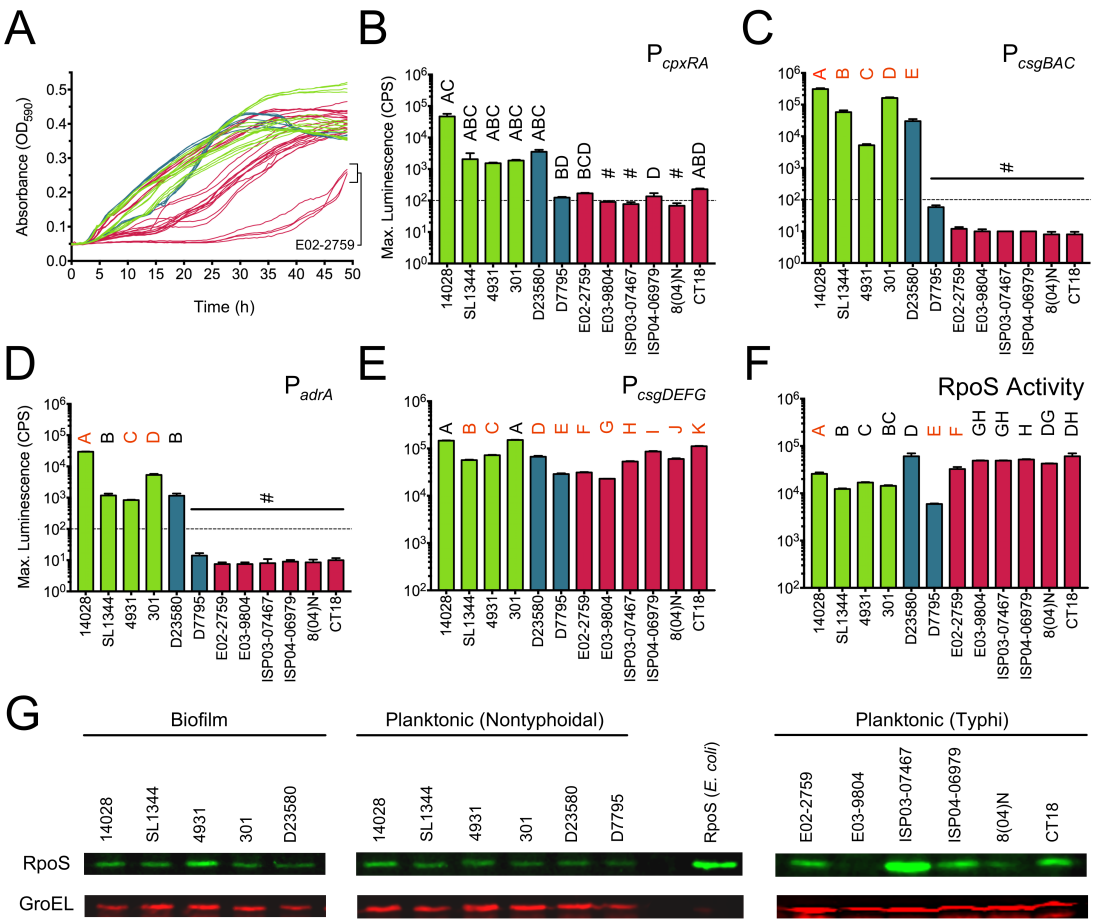

Figure 5.2. Analysis of the biofilm regulatory network in nontyphoidal and typhoidal *Salmonella* strains.

(A) Absorbance measurements representing strain growth over 48 hours at 28˚C in microaerophilic cultures grown in 96-well microtiter plates. (A-F) Green - nontyphoidal *Salmonella* strains associated with gastroenteritis disease; blue - nontyphoidal *Salmonella* strains associated with invasive bloodstream infections; and pink - typhoidal *Salmonella* strains. (B-E) Maximal luminescence (counts per second, CPS) is reported for *Salmonella* strains harbouring vectors containing *S.* Typhimurium 14028 promoters fused to bacterial luciferase genes (*luxCDABE*). (F) Maximal luminescence from strains containing a vector-based synthetic RpoS-responsive promoter-luciferase fusion construct indicating the functional activity of the RpoS sigma factor in each strain. The maximal luminescence value reported represents the mean and standard deviation of values recorded from a minimum of 3 independent biological replicates. We used an ordinary

one-way ANOVA with Holm-Sidak's multiple comparison test or a Kruskal-Wallis test with Dunn's multiple comparisons test to detect strain-dependent variations in the maximal luciferase expression of a given promoter-reporter construct. Mean values that are not statistically different from other mean values are indicated with a common letter in black font. In contrast, mean values that are statistically different from all other mean values are indicated with a unique letter in red font. The dashed line in (B), (C), and (D) represents the limit of detection for promoter expression and is derived from the maximal luminescence values measured from *S.* Typhimurium 14028 containing a promoterless vector construct; #, measured value below limit of detection. (G) Western blot analysis of RpoS sigma factor protein in whole cell lysates of multicellular aggregates or planktonic cells derived from nontyphoidal and typhoidal *Salmonella* strains collected following 24 hours of growth at 28˚C in liquid cultures of 1% tryptone. Recombinant RpoS protein purified from *E. coli* cells was used as a technical control for RpoS detection. GroEL was detected as a loading control to ensure that each lane was loaded with equal amounts of protein. The Western blots are representative of two biological replicates.

29,000 CPS in biofilm-positive strains and was strongest in *S.* Typhimurium 14028 (Fig. 5.2D). In contrast, expression of *csgBAC* and *adrA* in *S.* Enteritidis D7795 and all *S.* Typhi strains was comparable to reporter noise recorded from control strains containing a promoterless reporter vector.

The *csgD* promoter region is targeted by several transcriptional regulators and represents an important site of integration for environmental signals, such as temperature, osmolarity, and nutritional limitation. We reasoned that loss of the biofilm phenotype in *S.* Enteritidis D7795 and *S.* Typhi strains could result from mutations affecting *trans*-acting regulatory elements that modulate the expression of *csgD*. We detected *csgD* promoter activity in all of the *Salmonella* strains included in our panel (Fig. 5.2E). However, *S.* Enteritidis D7795, *S.* Typhi E02-2759, and *S.* Typhi E03-9804 demonstrated a 5- to 6-fold reduction in maximal promoter expression compared to *S.* Typhimurium 14028. For these strains, loss of the biofilm phenotype may reflect the presence of mutations in the regulatory network that consequentially result in reduced *csgD* promoter expression.

Expression of *csgD* is dependent on RpoS, a sigma factor with maximal expression during stationary phase that promotes the transcription of genes important for cell survival under conditions of starvation or stress (35). While *rpoS* mutant alleles are common in *S.* Typhi, they are relatively rare in nontyphoidal *Salmonella* isolates and are often associated with strain passaging and growth in laboratory conditions (32, 60-62). We tracked RpoS activity in our panel of strains by using an RpoS-responsive synthetic promoter fused to the *luxCDABE* operon, sig38H4 (26). This promoter was previously demonstrated to have a maximal activity of less than 10,000 CPS in *rpoS* deficient strains of *Salmonella* (32). The majority of strains in our panel exceeded this threshold; however, maximal expression of the *sig38H4* promoter in *S.* Enteritidis D7795 was significantly different from the expression levels observed for all other strains in the panel, with a mean value of 5,900 CPS (Fig. 5.2F). To confirm the presence of RpoS at the protein level, we probed whole-cell lysates of each strain obtained from the flask cultures we had used to assess for the biofilm phenotype (Fig. 5.1). The sigma factor was detectable in all nontyphoidal *Salmonella* strains, including both multicellular aggregate and planktonic cell subpopulations derived from biofilm-producing strains (Fig. 5.2G).

We noted significant variation in RpoS protein levels detected in each strain at the 24-hour time point used for sampling, with nearly undetectable levels of RpoS observed for E03-9804 and 8(04)N.

5.5.3 Species and strain-specific allelic variation in *csg* promoter expression and protein synthesis

We hypothesized that mutations affecting *csgD* expression or synthesis may account for the significant reduction in *csgBAC* promoter expression for *S.* Typhimurium D23580 or absence of *csgBAC* and *adrA* promoter expression for *S.* Enteritidis D7795. To identify these potential mutations, we performed a genetic alignment of the sequences for the *csg* region from the nontyphoidal and typhoidal strains included in our panel (Fig. 5.3A). Sequence identity was high overall, scoring 98% for the full-length *csg* region (4,450 bp), and 94% for the intergenic region (755 bp). Several serovar-specific sequence changes were observed for both *csg* operons as well as the intergenic region. Our alignment did not identify any mutations within the *csg* operons for *S.* Typhimurium D23580 or *S.* Enteritidis D7795 (data not shown). However, we identified unique single nucleotide polymorphisms within the intergenic region that were specific to either strain of invasive nontyphoidal *Salmonella* (Fig. 5.3B). For *S.* Typhimurium D23580, a G-to-T transversion was found 80 bp upstream of the *csgD* transcriptional start site, where OmpR and H-NS have been demonstrated to bind and subsequently repress *csgD* transcription (63). A second mutation in the *S.* Typhimurium D23580 sequence, a C-to-T transition, was observed 189 bp upstream. This area of the intergenic region is linked to both enhancement and repression of *csgD* transcription, resulting from an overlap of an activating binding site for integration host factor (IHF) and inactivating binding sites for OmpR and the nucleoid associated factor H-NS (63, 64). For *S.* Enteritidis D7795, a G-to-A transition mutation was observed at position 47 upstream of the *csgD* transcriptional start site, an area corresponding to a high-affinity binding site for OmpR and an important region for the transcriptional activation of *csgD* (63).

To determine the impact of these polymorphisms, we amplified the *csg* intergenic region to generate a set of *csgDEFG* and *csgBAC* promoter-luciferase fusion vectors from each strain. We then compared the functionality of each promoter by measuring their

expression levels in our biofilm-positive control strain, *S.* Typhimurium 14028 (Fig. 5.3C). We observed a similar level of maximal expression for the *csgD* promoters from our control strains of nontyphoidal and typhoidal *Salmonella*. Similarly, peak expression measured for the *csgD* promoters for *S.* Typhimurium D23580 and *S.* Enteritidis D7795 was comparable to most of the strains included in our panel. However, our statistical analysis revealed a significant difference in *csgD* promoter expression between either invasive nontyphoidal *Salmonella* strain and *S.* Typhimurium 14028. In agreements with our multiple sequence alignments, we found that the functionality of the *S.* Typhimurium D23580 and *S.* Enteritidis D7795 *csgB* promoters was not statistically different from that of either nontyphoidal or typhoidal promoters. Surprisingly, expression from the *csgB* promoter of biofilm-producing strain *S.* Typhimurium SL1344 was statistically below expression levels for promoters from half of the strains in our panel.

We hypothesized that variation in *csgD* promoter expression may account for differences in the biofilm phenotype between both *S.* Typhimurium D23580 and *S.* Enteritidis D7795 and nontyphoidal *Salmonella* strains associated with gastroenteritis or invasive disease. Using western blot, we evaluated CsgD synthesis and curli fimbriae biogenesis in whole cell lysates obtained from multicellular aggregates for biofilm-positive strains and homogeneous planktonic cultures of biofilm-negative strains (Fig. 5.3D). Consistent with our previous work (45), both CsgD protein and CsgA, the major subunit of curli fimbriae, was detected in lysates from multicellular aggregates for all biofilm-positive strains, including *S.* Typhimurium D23580. In contrast, CsgD and CsgA synthesis was not observed for *S.* Enteritidis D7795 or any of the *S.* Typhi strains. While *csgD* promoter strength was comparable between *S.* Typhi and our control strains of biofilm-positive nontyphoidal *Salmonella*, absence of CsgD synthesis may result from predicted truncation of 8 amino acids from the N-terminal end that is conserved across all typhoidal strains in our panel (14, 16).

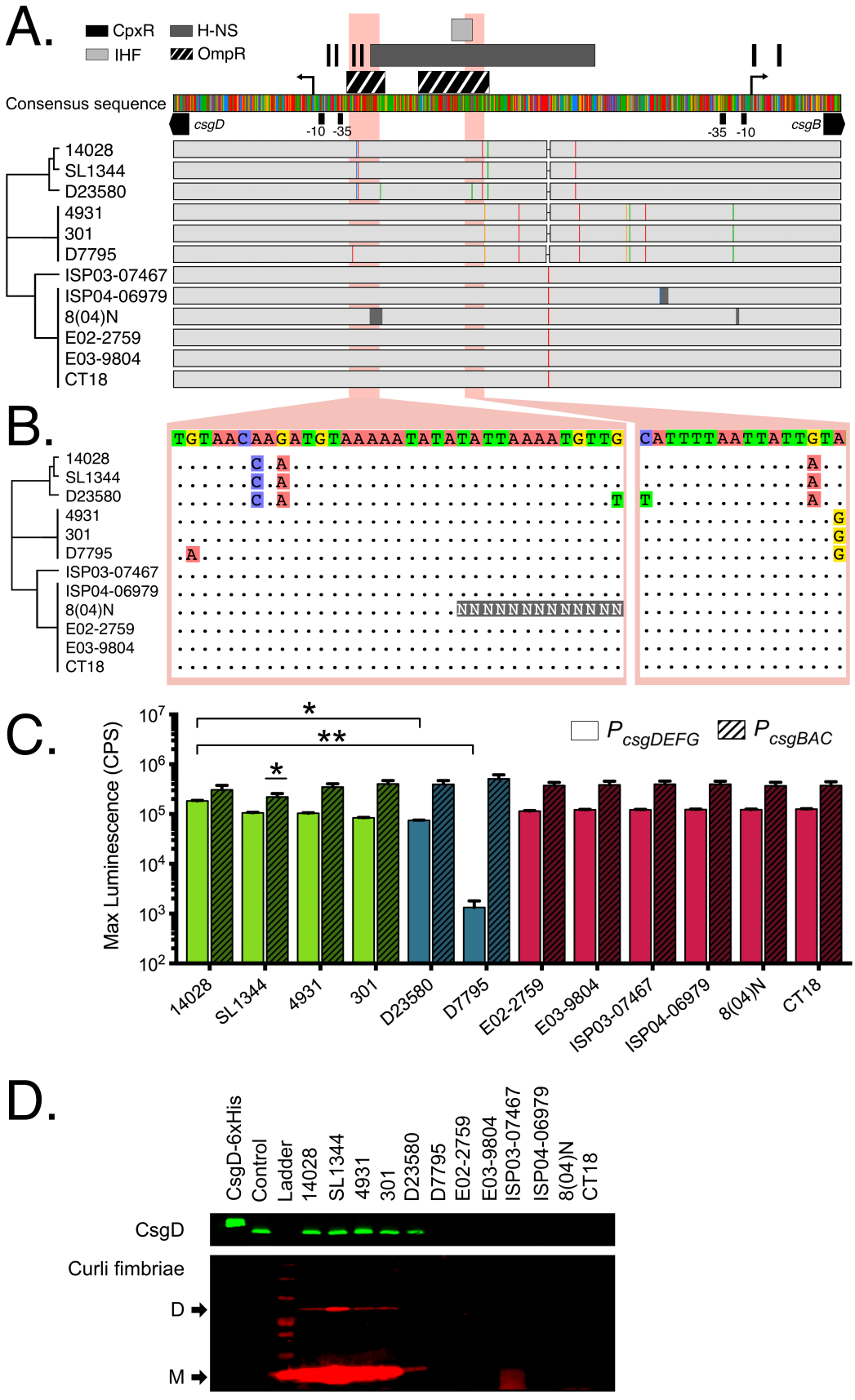
A.
CpxR
IHF
H-NS
OmpR
Consensus sequence
csgD
-10
-35
-35
-10
csgB
14028
SL1344
D23580
4931
301
D7795
ISP03-07467
ISP04-06979
8(04)N
E02-2759
E03-9804
CT18
B.
TGTAACAAGATGTAAAAATATATATTAAAATGTTG
CATTTTAATTATTGTA
NNNNNNNNNNNNNN
C.
Max Luminescence (CPS)
10^7
10^6
10^5
10^4
10^3
10^2
*
**
*
$P_{csgDEFG}$
P_{csgBAC}
14028
SL1344
4931
301
D23580
D7795
E02-2759
E03-9804
ISP03-07467
ISP04-06979
8(04)N
CT18
D.
CsgD-6xHis
Control
Ladder
14028
SL1344
4931
301
D23580
D7795
E02-2759
E03-9804
ISP03-07467
ISP04-06979
8(04)N
CT18
CsgD
Curli fimbriae
D
M

Figure 5.3. Identification of sequence changes and comparison of promoter activity in the *csgD-csgB* intergenic region from different *Salmonella* strains.

(A) Multiple sequence alignment of the intergenic and 5′ untranslated regions for the *csgDEFG* and *csgBAC* operons from each *Salmonella* strain included in this study. The neighbour-joining dendrograms included to the left of the sequence alignments are established based on bootstrapping parameters set to 1,000 replicates and a support threshold of 70%. Both transcriptional start sites (arrows) and transcription factor binding regions that have been experimentally verified in *Salmonella* are indicated above the consensus sequence. CpxR, black boxes; H-NS, dark grey boxes; IHF, light grey boxes; OmpR, hatched boxes. The start of the *csgD* and *csgB* open reading frames and the -10 or -35 regions for either operon are included below the consensus sequence. Red boxes that span the sequence alignment indicate regions containing strain-specific mutations for *S.* Typhimurium D23850 or *S.* Enteritidis D7795, which are expanded in (B). (C) Maximum luminescence from promoter-reporter constructs driven by *csgD* and *csgB* promoters from each *Salmonella* strain. Expression of each reporter was measured in *S.* Typhimurium 14028. The maximal luminescence value reported represents the mean and standard deviation of values recorded from 3 independent biological replicates. Statistical significance: *, $P < 0.05$; **, $P < 0.01$. (D) Whole cell lysates of multicellular aggregates or planktonic cells from biofilm-positive or biofilm-negative strains, respectively, were analyzed by western blotting for the presence of CsgD and CsgA, the major subunit of curli fimbriae. Whole cell lysates were derived from nontyphoidal and typhoidal *Salmonella* strains collected following 24 hours of growth at 28˚C in liquid cultures of 1% tryptone. Black arrows indicate monomerized (M) and dimerized (D) CsgA protein subunits. A CsgD-6xHis recombinant protein was used as a technical control for CsgD detection. The control lane was used for inter-blot normalization of band intensities and represents the pooled whole cell lysates derived from *S.* Typhimurium strains 14028 and SL1344, as well as *S.* Enteritidis 4931 and 301. The western blots are representative of two biological replicates.

5.5.4 A missense mutation in the Cache1 domain of STM1987 results in differential expression of cellulose between *S.* Typhimurium D23580 and 14028

BcsA, the catalytic subunit for cellulose biosynthesis, is allosterically activated via binding of the bacterial secondary messenger molecule bis-(3',5')-cyclic dimeric guanosine monophosphate (c-di-GMP) to the protein's PilZ domain (65). While AdrA is the predominant contributor to this inducing c-di-GMP pool during CsgD-regulated biofilm formation, additional diguanylate cyclases such as STM1987 further enhance intracellular levels of this secondary messenger and promote cellulose production (45, 66-68). Mills *et al.* recently identified the Cache1 domain of the STM1987 enzyme as the potential activation site for enzyme activity in response to inducing molecules such as L-arginine (69). We compared the gene sequence for STM1987 from *S.* Typhimurium strains 14028 and D23580 and observed a C-to-G transversion located 566 bp downstream of the translational start site. For *S.* Typhimurium D23580, this transversion results in a replacement of threonine with arginine at residue 189, located within the Cache1 domain region for STM1987.

To determine the effect of this mutation, the alleles for STM1987 from *S.* Typhimurium 14028 and D23580 were cloned into plasmid vectors and transformed into an *S.* Typhimurium Δ*STM1987* mutant strain (Fig. 5.4). Plasmid-based expression of diguanylate cyclases has been previously demonstrated to result in the overexpression of cellulose in otherwise non-inducing conditions (66, 70). Consistent with these observations, either complemented strain demonstrated uncharacteristic biofilm production at the air-liquid interface following overnight growth in LB broth at 37˚C (Fig. 5.4A). We noted a difference in the thickness of this biofilm layer between cultures of each complemented strain, suggesting that the STM1987 allele from *S.* Typhimurium D23580 may have decreased enzymatic activity. This effect was visible during macrocolony growth of either complemented strain on T agar containing calcofluor white (Fig. 5.4B). Following 12 hours of growth at 28˚C, the fluorescent intensity of macrocolonies of the STM1987 (D23580) complemented strain was observably less than that of the STM1987 (14028) complemented strain (Fig. 5.4C). However, these differences in fluorescence were no longer visible by 20 hours of growth (Fig. 5.4D).

5.6 Discussion

In this study, we evaluated the biofilm phenotype of invasive nontyphoidal strains *S.* Typhimurium D23580 and *S.* Enteritidis D7795. Our results showed that invasive nontyphoidal Salmonellae vary in their capacity for biofilm formation. While *S.* Typhimurium D23580 demonstrated a partial rdar morphotype and produced multicellular aggregates in liquid cultures, *S.* Enteritidis D7795 was rdar-negative and solely consisted of a homogenous planktonic population. Consistent with these phenotypic differences, we observed that curli fimbriae and cellulose extracellular matrix components could be detected from macrocolonies or multicellular aggregates of *S.* Typhimurium D23580, but not *S.* Enteritidis D7795. Our sequence alignment of *csgD* promoter regions from multiple nontyphoidal and typhoidal *Salmonella* strains revealed a small nucleotide polymorphism unique to *S.* Enteritidis D7795, located in a critical OmpR binding site required for activation of *csgD* transcriptional expression (63). This mutation resulted in a drop in *csgD* promoter activity to near background levels, which correlates with the absence of *csgBAC* and *adrA* promoter expression in this invasive *S.* Enteritidis strain.

Our findings contradict a recent study from Singletary *et al.* that had demonstrated a rdar-negative phenotype for *S.* Typhimurium D23580 (43). Their comparative analysis of the *csg* and *bcs* operons of *S.* Typhimurium strains 14028 and D23580 identified three potential mutations, including a nonsynonymous mutation within *csgE* and a premature stop codon in *bcsG*. Our multiple sequence alignment of the *csg* operons from strains included in this study found that this allelic difference in *csgE*, located 108 bp from the *csgE* translational start site, is in fact unique to *S.* Typhimurium 14028 compared to all other nontyphoidal *Salmonella* strains in this study (data not shown). Singletary *et al.* demonstrated that complementation of *S.* Typhimurium D23580 with a plasmid-based copy of the *bcsG* allele from *S.* Typhimurium 14028 resulted in rescue of the rdar morphotype. BcsG encodes a putative endoglucanase, an enzyme that

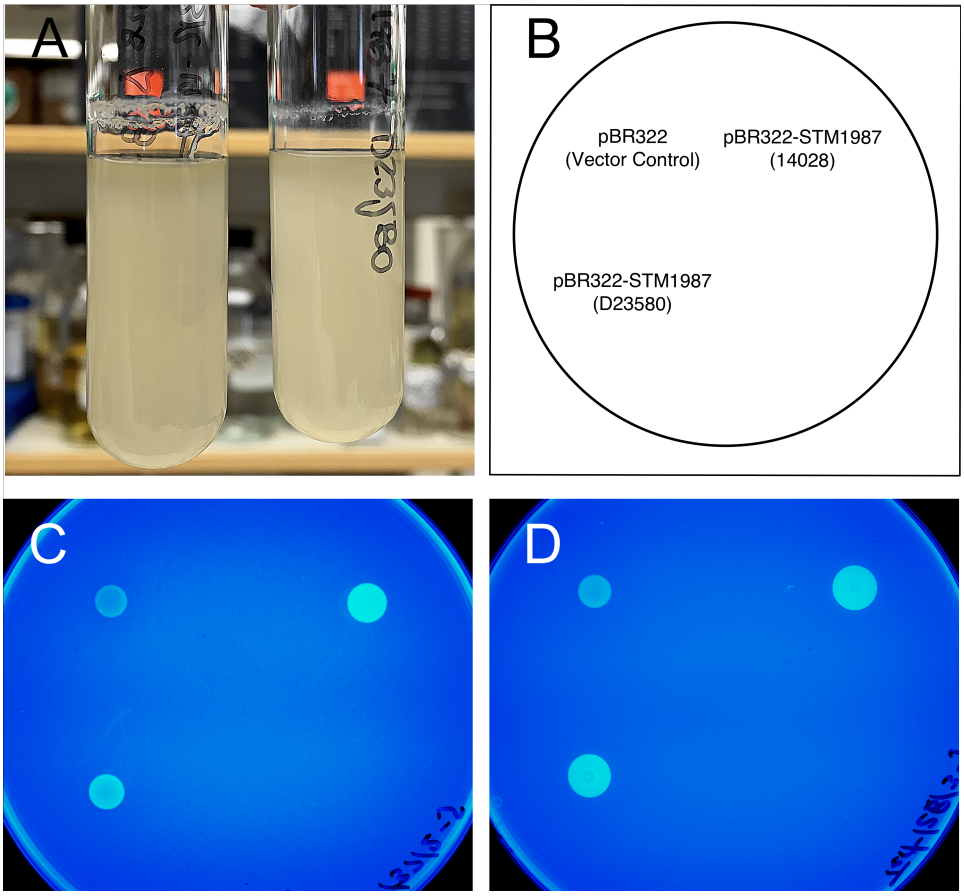

Figure 5.4. Phenotypic comparison between STM1987 alleles from *S.* Typhimurium 14028 and D23580.

(A) Biofilm production was visible as an adhered ring of cells and extracellular matrix production in overnight LB broth cultures of *S.* Typhimurium Δ*STM1987* complemented by vector-based alleles of STM1987 from *S.* Typhimurium 14028 or D23580. (B) To observe the effect of STM1987 complementation on cellulose production, we spotted 2 µL of each overnight culture on T agar supplemented with 200 µg mL^{-1} calcofluor white dye. *S.* Typhimurium 14028 Δ*STM1987* containing an empty pBR322 vector was used as a negative control. Macrocolony binding of calcofluor white was visualized by UV light after (C) 12 and (D) 20 hours of growth at 28˚C.

is normally responsible for cleavage of polymeric cellulose (34, 71). While the *bcsEFG* operon has been shown as necessary for optimal cellulose production in *S.* Enteritidis, the exact role and level of importance of BcsG has yet to be determined (34, 72). We demonstrated that macrocolonies of *S.* Typhimurium D23580 retain the ability to produce cellulose despite this mutated *bcsG* allele. However, mutation of the *bcsG* gene and sub-optimal cellulose biosynthesis may represent one of many factors that contribute to the impaired biofilm phenotype that we observed for *S.* Typhimurium D23580.

We identified a missense mutation in the *S.* Typhimurium D23580 sequence for STM1987, a diguanylate cyclase responsible for c-di-GMP biosynthesis. This important secondary messenger molecule is involved in the regulation of multiple cellular processes for enteric bacteria, including motility, biofilm formation, and virulence (reviewed in (73)). In *Salmonella*, at least 5 diguanylate cyclases and 4 phosphodiesterases have been associated with modulation of the c-di-GMP pool that regulates the rdar morphotype (50, 66-68, 73-76). However, STM1987 is the only other diguanylate cyclase apart from AdrA that induces cellulose biosynthesis under physiological conditions, without the requirement for overexpression (45, 67). Our improved understanding of enzyme-specific regulation, including the conditions for transcriptional expression, membrane versus cytoplasmic enzyme localization, and signals that induce enzymatic activity, may explain how c-di-GMP enzymes can produce the same messenger molecule but affect different levels in the biofilm regulatory pathway (73). While transcription of both STM1987 and AdrA is maximal during stationary phase, STM1987 is unique in that its transcription is not dependent on CsgD synthesis (N. Herman and A. White, unpublished data, (77)). Pontes *et al.* recently provided evidence for cellulose production by *S.* Typhimurium while within the *Salmonella*-containing vacuole of infected macrophages (78). STM1987 may induce biosynthesis of cellulose under these conditions in response to sensing increased vacuolar concentrations of L-arginine via its Cache1 domain (69). It is tempting to speculate that decreased activity of STM1987 in *S.* Typhimurium D23580 may result from differences in macrophage activation and intracellular killing in immunocompromised individuals. Reduced cellulose biosynthesis during infection may improve the virulence and replication of *S.* Typhimurium D23580 in its

immunocompromised host, but would impair its ability to produce the rdar morphotype and survive in the environment between host infections (12, 38, 45).

Both *csgD* expression and the rdar morphotype are highly conserved across the *Salmonella* species (32, 34, 63, 79). Strains that are negative for the rdar morphotype, including *Salmonella* Typhi, Paratyphi A, Choleraesuis, Gallinarum, and Typhimurium var. Copenhagen, are often associated with a restricted host range and are responsible for invasive forms of salmonellosis in their respective hosts (2, 31, 32). The transition from the intestinal to systemic niche is thought to represent an evolutionary bottleneck for *Salmonella*, with significant losses in the functional gene repertoire consistently observed for invasive serovars and strains of *Salmonella* (12, 80, 81). Some serovars and strains, including *S.* Typhi, *S.* Paratyphi A, as well as invasive nontyphoidal *Salmonella*, have demonstrated an ability to develop an asymptomatic chronic carrier state within the human host, removing the selection pressure on genes that may specifically aid in transmission via the fecal-oral route (37, 38). As the CsgD-regulated biofilm phenotype is hypothesized to aid cell survival during the environmental phase of transmission, it is possible that genes for this phenotype are functionally inactivated as part of host adaptation (25, 26, 31, 34, 45, 82, 83). However, the function of some biofilm-related processes, such as exopolysaccharide production, may be selected due to their importance for CsgD-independent biofilm formation in the gall bladder (8, 84, 85).

We recently reported the increased expression of over 780 genes during the development of multicellular aggregates in cultures of *S.* Typhimurium 14028 (45). This significant shift in the transcriptome of *Salmonella* cells contrasts sharply with both the small regulon of genes directly controlled by CsgD as well as the complex but limited number of regulatory factors that influence expression and synthesis of CsgD itself ((86), reviewed in (87)). It is possible that patterns of genomic degradation that accompany host adaptation may affect the biofilm phenotype in a graded manner. Neutral or mildly deleterious mutations affecting the core CsgD regulatory pathway or functional inactivation of genes included in the transcriptome may result in impairment of the rdar morphotype. However, gene inactivation or deletions in the core biofilm regulatory pathway would result in loss of the rdar morphotype altogether. It is interesting to note that variation in the biofilm phenotype of invasive nontyphoidal *Salmonella* strains

included in this study reflects differences in the known host range of either strain. *S.* Typhimurium D23580 demonstrated a partial rdar morphotype and has been shown to cause infection in several host species, including humans, mice, chicks, cows, and rhesus macaques (43). In contrast, *S.* Enteritidis D7795 was negative for the rdar morphotype, and has thus far been demonstrated to have reduced invasiveness and overall fitness during experimental infection of chicks (19). We propose that genetic changes impacting the CsgD-regulated biofilm phenotype result in the loss of an environmental persistence strategy that is central to ancestral strains of *E. coli* and *Salmonella* that transmit via the fecal-oral route, and may therefore represent an important genetic signature of host adaptation.

5.7 Acknowledgements

This research was supported by a grant from the Natural Sciences and Engineering Research Council (NSERC) (#386063-2010) to APW and through the Jarislowsky Chair in Biotechnology. KDM was supported by a Canada Graduate Scholarship from NSERC, and DJH was supported by an Undergraduate Research Award from the University of Saskatchewan. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

The authors are grateful to Michael Prodanuk and Melissa Palmer for laboratory assistance; and to Jay Hinton, Andrew Cameron, and Alex Pasternak for helpful discussions.

5.8 References

1. **Grimont P**, **Weill FX**. 2007. Antigenic formulae of the *Salmonella* serotypes. France: WHO Collaborating Centre for Reference and Research on *Salmonella*. **9**.

2. **Waldner L**, **MacKenzie K**, **Köster W**, **White A**. 2012. From Exit to Entry: Long-term Survival and Transmission of *Salmonella*. Pathogens 2012, Vol 1, Pages 128-155 **1**:128–155.

3. **Kirk MD**, **Pires SM**, **Black RE**, **Caipo M**, **Crump JA**, **Devleesschauwer B**, **Döpfer D**, **Fazil A**, **Fischer-Walker CL**, **Hald T**, **Hall AJ**, **Keddy KH**, **Lake RJ**, **Lanata CF**, **Torgerson PR**, **Havelaar AH**, **Angulo FJ**. 2015. World Health Organization Estimates of the Global and Regional Disease Burden of 22

Foodborne Bacterial, Protozoal, and Viral Diseases, 2010: A Data Synthesis. PLoS Med **12**:e1001921–21.

4. **Majowicz SE**, **Musto J**, **Scallan E**, **Angulo FJ**, **Kirk M**, **O'Brien SJ**, **Jones TF**, **Fazil A**, **Hoekstra RM**. 2010. The Global Burden of Nontyphoidal *Salmonella* Gastroenteritis. Clinical Infectious Diseases **50**:882–889.

5. **Buckle GC**, **Walker CLF**, **Black RE**. 2012. Typhoid fever and paratyphoid fever: Systematic review to estimate global morbidity and mortality for 2010. J Glob Health **2**:010401.

6. **Dougan G**, **Baker S**. 2014. *Salmonella enterica* Serovar Typhi and the Pathogenesis of Typhoid Fever. Annu Rev Microbiol **68**:317–336.

7. **Parry CM**, **Hien TT**, **Dougan G**, **White NJ**, **Farrar JJ**. 2002. Typhoid fever. N Engl J Med **347**:1770–1782.

8. **Gunn JS**, **Marshall JM**, **Baker S**, **Dongol S**, **Charles RC**, **Ryan ET**. 2014. *Salmonella* chronic carriage: epidemiology, diagnosis, and gallbladder persistence. Trends Microbiol **22**:648–655.

9. **Reddy EA**, **Shaw AV**, **Crump JA**. 2010. Community-acquired bloodstream infections in Africa: a systematic review and meta-analysis. The Lancet Infectious Diseases **10**:417–432.

10. **Feasey NA**, **Dougan G**, **Kingsley RA**, **Heyderman RS**, **Gordon MA MD**. 2012. Invasive non-typhoidal *Salmonella* disease: an emerging and neglected tropical disease in Africa. The Lancet **379**:2489–2499.

11. **Gordon MA**. 2008. *Salmonella* infections in immunocompromised adults. Journal of Infection **56**:413–422.

12. **Baumler A**, **Fang FC**. 2013. Host Specificity of Bacterial Pathogens. Cold Spring Harbor Perspectives in Medicine **3**:a010041–a010041.

13. **Betancor L**, **Yim L**, **Martínez A**, **Fookes M**, **Sasias S**, **Schelotto F**, **Thomson N**, **Maskell D**, **Chabalgoity JA**. 2012. Genomic Comparison of the Closely Related *Salmonella enterica* Serovars Enteritidis and Dublin. Open Microbiol J **6**:5–13.

14. **Parkhill J**, **Dougan G**, **James KD**, **Thomson NR**, **Pickard D**, **Wain J**, **Churcher C**, **Mungall KL**, **Bentley SD**, **Holden MT**, **Sebaihia M**, **Baker S**, **Basham D**, **Brooks K**, **Chillingworth T**, **Connerton P**, **Cronin A**, **Davis P**, **Davies RM**, **Dowd L**, **White N**, **Farrar J**, **Feltwell T**, **Hamlin N**, **Haque A**, **Hien TT**, **Holroyd S**, **Jagels K**, **Krogh A**, **Larsen TS**, **Leather S**, **Moule S**, **O'Gaora P**, **Parry C**, **Quail M**, **Rutherford K**, **Simmonds M**, **Skelton J**, **Stevens K**, **Whitehead S**, **Barrell BG**. 2001. Complete genome sequence of a multiple drug resistant *Salmonella enterica* serovar Typhi CT18. Nature **413**:848–852.

15. **McClelland M**, **Sanderson KE**, **Clifton SW**, **Latreille P**, **Porwollik S**, **Sabo A**, **Meyer R**, **Bieri T**, **Ozersky P**, **McLellan M**, **Harkins CR**, **Wang C**, **Nguyen C**, **Berghoff A**, **Elliott G**, **Kohlberg S**, **Strong C**, **Du F**, **Carter J**, **Kremizki C**, **Layman D**, **Leonard S**, **Sun H**, **Fulton L**, **Nash W**, **Miner T**, **Minx P**, **Delehaunty K**, **Fronick C**, **Magrini V**, **Nhan M**, **Warren W**, **Florea L**, **Spieth J**, **Wilson RK**. 2004. Comparison of genome degradation in Paratyphi A and Typhi, human-restricted serovars of *Salmonella enterica* that cause typhoid. Nature Genetics **36**:1268–1274.

16. **Holt KE**, **Parkhill J**, **Mazzoni CJ**, **Roumagnac P**, **Weill F-X**, **Goodhead I**, **Rance R**, **Baker S**, **Maskell DJ**, **Wain J**, **Dolecek C**, **Achtman M**, **Dougan G**. 2008. High-throughput sequencing provides insights into genome variation and evolution in *Salmonella* Typhi. Nature Genetics **40**:987–993.

17. **Kingsley RA**, **Msefula CL**, **Thomson NR**, **Kariuki S**, **Holt KE**, **Gordon MA**, **Harris D**, **Clarke L**, **Whitehead S**, **Sangal V**, **Marsh K**, **Achtman M**, **Molyneux ME**, **Cormican M**, **Parkhill J**, **MacLennan CA**, **Heyderman RS**, **Dougan G**. 2009. Epidemic multiple drug resistant *Salmonella* Typhimurium causing invasive disease in sub-Saharan Africa have a distinct genotype. Genome Research **19**:2279–2287.

18. **Okoro CK**, **Barquist L**, **Connor TR**, **Harris SR**, **Clare S**, **Stevens MP**, **Arends MJ**, **Hale C**, **Kane L**, **Pickard DJ**, **Hill J**, **Harcourt K**, **Parkhill J**, **Dougan G**, **Kingsley RA**. 2015. Signatures of Adaptation in Human Invasive *Salmonella* Typhimurium ST313 Populations from Sub-Saharan Africa. PLoS Negl Trop Dis **9**:e0003611–18.

19. **Hadfield J**, **Keddy KH**, **Dallman TJ**, **Jacobs J**, **Deng X**, **Wigley P**, **Barquist LB**, **Langridge GC**, **Feltwell T**, **Harris SR**, **Mather AE**, **Fookes M**, **Aslett M**, **Msefula C**, **Kariuki S**, **MacLennan CA**, **Onsare RS**, **Weill FCO-X**, **Le Hello S**, **Smith AM**, **McClelland M**, **Desai P**, **Parry CM**, **Cheesbrough J**, **French N**, **Campos J**, **Chabalgoity JA**, **Betancor L**, **Hopkins KL**, **Nair S**, **Humphrey TJ**, **Lunguya O**, **Cogan TA**, **Tapia MD**, **Sow SO**, **Tennant SM**, **Bornstein K**, **Levine MM**, **Lacharme-Lora L**, **Everett DB**, **Kingsley RA**, **Parkhill J**, **Heyderman RS**, **Dougan G**, **Gordon MA**, **Thomson NR**, **Feasey NA**. 2016. Distinct *Salmonella* Enteritidis lineages associated with enterocolitis in high-income settings and invasive disease in low-income settings. Nature Genetics 1–9.

20. **Winter SE**, **Thiennimitr P**, **Winter MG**, **Butler BP**, **Huseby DL**, **Crawford RW**, **Russell JM**, **Bevins CL**, **Adams LG**, **Tsolis RM**, **Roth JR**, **Bäumler AJ**. 2010. Gut inflammation provides a respiratory electron acceptor for *Salmonella*. Nature **467**:426–429.

21. **Thiennimitr P**, **Winter SE**, **Winter MG**, **Xavier MN**, **Tolstikov V**, **Huseby DL**, **Sterzenbach T**, **Tsolis RM**, **Roth JR**, **Bäumler AJ**. 2011. Intestinal inflammation allows *Salmonella* to use ethanolamine to compete with the microbiota. Proceedings of the National Academy of Sciences of the United States of America

108:17480–17485.

22. **Nuccio SP**, **Baumler AJ**. 2014. Comparative Analysis of *Salmonella* Genomes Identifies a Metabolic Network for Escalating Growth in the Inflamed Gut. mBio **5**:e00929–14–e00929–14.

23. **Lawley TD**, **Bouley DM**, **Hoy YE**, **Gerke C**, **Relman DA**, **Monack DM**. 2008. Host transmission of *Salmonella enterica* serovar Typhimurium is controlled by virulence factors and indigenous intestinal microbiota. Infection and Immunity **76**:403–416.

24. **Winfield MD**, **Groisman EA**. 2003. Role of Nonhost Environments in the Lifestyles of *Salmonella* and *Escherichia coli*. Applied and Environmental Microbiology **69**:3687–3694.

25. **Scher K**, **Romling U**, **Yaron S**. 2005. Effect of Heat, Acidification, and Chlorination on *Salmonella enterica* Serovar Typhimurium Cells in a Biofilm Formed at the Air-Liquid Interface. Applied and Environmental Microbiology **71**:1163–1168.

26. **White AP**, **Gibson DL**, **Kim W**, **Kay WW**, **Surette MG**. 2006. Thin aggregative fimbriae and cellulose enhance long-term survival and persistence of *Salmonella*. Journal of Bacteriology **188**:3219–3227.

27. **Collinson SK**, **Emödy L**, **Müller KH**, **Trust TJ**, **Kay WW**. 1991. Purification and characterization of thin, aggregative fimbriae from *Salmonella enteritidis*. Journal of Bacteriology **173**:4773–4781.

28. **Romling U**, **Sierralta WD**, **Eriksson K**, **Normark S**. 1998. Multicellular and aggregative behaviour of *Salmonella typhimurium* strains is controlled by mutations in the agfD promoter. Molecular Microbiology **28**:249–264.

29. **Zogaj X**, **Nimtz M**, **Rohde M**, **Bokranz W**, **Romling U**. 2001. The multicellular morphotypes of *Salmonella typhimurium* and *Escherichia coli* produce cellulose as the second component of the extracellular matrix. Molecular Microbiology **39**:1452–1463.

30. **Gibson D**, **White A**, **Snyder S**, **Martin S**, **Heiss C**, **Azadi P**, **Surette M**, **Kay W**. 2006. *Salmonella* produces an O-antigen capsule regulated by AgfD and important for environmental persistence. Journal of Bacteriology **188**:7722.

31. **Römling U**, **Bokranz W**, **Rabsch W**, **Zogaj X**, **Nimtz M**, **Tschäpe H**. 2003. Occurrence and regulation of the multicellular morphotype in *Salmonella* serovars important in human disease. International Journal of Medical Microbiology **293**:273–285.

32. **White A**, **Surette M**. 2006. Comparative genetics of the rdar morphotype in *Salmonella*. Journal of Bacteriology **188**:8395.

33. **Doran JL**, **Collinson SK**, **Burian J**, **Sarlós G**, **Todd EC**, **Munro CK**, **Kay CM**, **Banser PA**, **Peterkin PI**, **Kay WW**. 1993. DNA-based diagnostic tests for *Salmonella* species targeting *agfA*, the structural gene for thin, aggregative fimbriae. Journal of Clinical Microbiology **31**:2263–2273.

34. **Solano C**, **García B**, **Valle J**, **Berasain C**, **Ghigo J-M**, **Gamazo C**, **Lasa I**. 2002. Genetic analysis of *Salmonella enteritidis* biofilm formation: critical role of cellulose. Molecular Microbiology **43**:793–808.

35. **Romling U**, **Bian Z**, **Hammar M**, **Sierralta WD**, **Normark S**. 1998. Curli fibers are highly conserved between *Salmonella typhimurium* and *Escherichia coli* with respect to operon structure and regulation. Journal of Bacteriology **180**:722–731.

36. **White AP**, **Sibley KA**, **Sibley CD**, **Wasmuth JD**, **Schaefer R**, **Surette MG**, **Edge TA**, **Neumann NF**. 2011. Intergenic Sequence Comparison of *Escherichia coli* Isolates Reveals Lifestyle Adaptations but Not Host Specificity. Applied and Environmental Microbiology **77**:7620–7632.

37. **Gonzalez-Escobedo G**, **Marshall JM**, **Gunn JS**. 2011. Chronic and acute infection of the gall bladder by *Salmonella* Typhi: understanding the carrier state. Nature Publishing Group **9**:9–14.

38. **Kariuki S**, **Revathi G**, **Kariuki N**, **Kiiru J**, **Mwituria J**, **Muyodi J**, **Githinji JW**, **Kagendo D**, **Munyalo A**, **Hart CA**. 2006. Invasive multidrug-resistant non-typhoidal *Salmonella* infections in Africa: zoonotic or anthroponotic transmission? Journal of Medical Microbiology **55**:585–591.

39. **Kariuki S**, **Revathi G**, **Gakuya F**, **Yamo V**, **Muyodi J**, **Hart CA**. 2002. Lack of clonal relationship between non-typhi *Salmonella* strain types from humans and those isolated from animals living in close contact. FEMS Immunol Med Microbiol **33**:165–171.

40. **Parsons BN**, **Humphrey S**, **Salisbury AM**, **Mikoleit J**, **Hinton JCD**, **Gordon MA**, **Wigley P**. 2013. Invasive non-typhoidal *Salmonella* Typhimurium ST313 are not host-restricted and have an invasive phenotype in experimentally infected chickens. PLoS Negl Trop Dis **7**:e2487.

41. **Yang J**, **Barrila J**, **Roland KL**, **Kilbourne J**, **Ott CM**, **Forsyth RJ**, **Nickerson CA**. 2015. Characterization of the Invasive, Multidrug Resistant Non-typhoidal *Salmonella* Strain D23580 in a Murine Model of Infection. PLoS Negl Trop Dis **9**:e0003839–17.

42. **Carden S**, **Okoro C**, **Dougan G**, **Monack D**. 2015. Non-typhoidal *Salmonella* Typhimurium ST313 isolates that cause bacteremia in humans stimulate less inflammasome activation than ST19 isolates associated with gastroenteritis. Pathogens and Disease **73**:ftu023–ftu023.

43. **Singletary LA**, **Karlinsey JE**, **Libby SJ**, **Mooney JP**, **Lokken KL**, **Tsolis RM**,

Byndloss MX, **Hirao LA**, **Gaulke CA**, **Crawford RW**, **Dandekar S**, **Kingsley RA**, **Msefula CL**, **Heyderman RS**, **Fang FC**. 2016. Loss of Multicellular Behavior in Epidemic African Nontyphoidal *Salmonella enterica* Serovar Typhimurium ST313 Strain D23580. mBio **7**:e02265–15–11.

44. **Bjarnason J**, **Southward CM**, **Surette MG**. 2003. Genomic profiling of iron-responsive genes in *Salmonella enterica* serovar Typhimurium by high-throughput screening of a random promoter library. Journal of Bacteriology **185**:4973–4982.

45. **MacKenzie KD**, **Wang Y**, **Shivak DJ**, **Wong CS**, **Hoffman LJL**, **Lam S**, **Kröger C**, **Cameron ADS**, **Townsend HGG**, **Köster W**, **White AP**. 2015. Bistable Expression of CsgD in *Salmonella enterica* Serovar Typhimurium Connects Virulence to Persistence. Infection and Immunity **83**:2312–2326.

46. **Kearse M**, **Moir R**, **Wilson A**, **Stones-Havas S**, **Cheung M**, **Sturrock S**, **Buxton S**, **Cooper A**, **Markowitz S**, **Duran C**, **Thierer T**, **Ashton B**, **Meintjes P**, **Drummond A**. 2012. Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics **28**:1647–1649.

47. **Datsenko KA**, **Wanner BL**. 2000. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. Proceedings of the National Academy of Sciences **97**:6640–6645.

48. **Maloy SR**, **Stewart VJ**, **Taylor RK**. 1996. Genetic analysis of pathogenic bacteria: a laboratory manual. Cold Spring Harbor Laboratory Press.

49. **Serra DO**, **Richter AM**, **Klauck G**, **Mika F**, **Hengge R**. 2013. Microanatomy at Cellular Resolution and Spatial Order of Physiological Differentiation in a Bacterial Biofilm. mBio **4**:e00103–13–e00103–13.

50. **Romling U**, **Rohde M**, **Olsén A**, **Normark S**, **Reinköster J**. 2000. AgfD, the checkpoint of multicellular and aggregative behaviour in *Salmonella typhimurium* regulates at least two independent pathways. Molecular Microbiology **36**:10–23.

51. **White A**, **Gibson D**, **Grassl G**, **Kay W**, **Finlay B**, **Vallance B**, **Surette M**. 2008. Aggregation via the red, dry, and rough morphotype is not a virulence adaptation in *Salmonella enterica* serovar Typhimurium. Infection and Immunity **76**:1048.

52. **Grantcharova N**, **Peters V**, **Monteiro C**, **Zakikhany K**, **Romling U**. 2010. Bistable expression of CsgD in biofilm development of *Salmonella enterica* serovar Typhimurium. Journal of Bacteriology **192**:456.

53. **Barnhart MM**, **Chapman MR**. 2006. Curli biogenesis and function. Annu Rev Microbiol **60**:131–147.

54. **Raivio TL**. 2014. Everything old is new again: An update on current research on the Cpx envelope stress response. BBA - Molecular Cell Research **1843**:1529–

1541.

55. **De Wulf P**, **Kwon O**, **Lin E**. 1999. The CpxRA Signal Transduction System of *Escherichia coli*: Growth-Related Autoactivation and Control of Unanticipated Target Operons. Journal of Bacteriology.

56. **Price NL**, **Raivio TL**. 2009. Characterization of the Cpx Regulon in *Escherichia coli* Strain MC4100. Journal of Bacteriology **191**:1798–1815.

57. **la Cruz De MA**, **Morgan JK**, **Ares MA**, **Yáñez-Santos JA**, **Riordan JT**, **Girón JA**. 2016. The Two-Component System CpxRA Negatively Regulates the Locus of Enterocyte Effacement of Enterohemorrhagic *Escherichia coli* Involving σ32 and Lon protease. Front Cell Infect Microbiol **6**:1985–13.

58. **Hammar M**, **Arnqvist A**, **Bian Z**, **Olsén A**, **Normark S**. 1995. Expression of two csg operons is required for production of fibronectin- and congo red-binding curli polymers in *Escherichia coli* K-12. Molecular Microbiology **18**:661–670.

59. **Zakikhany K**, **Harrington CR**, **Nimtz M**, **Hinton JCD**, **Römling U**. 2010. Unphosphorylated CsgD controls biofilm formation in *Salmonella enterica* serovar Typhimurium. Molecular Microbiology **77**:771–786.

60. **Jorgensen F**, **Leach S**, **Wilde SJ**, **Davies A**, **Stewart G**, **Humphrey T**. 2000. Invasiveness in chickens, stress resistance and RpoS status of wild-type *Salmonella enterica* subsp *enterica* serovar Typhimurium definitive type 104 and serovar Enteritidis phage type 4 strains. Microbiology **146**:3227–3235.

61. **Robbe-Saule V**, **Algorta G**, **Rouilhac I**, **Norel F**. 2003. Characterization of the RpoS Status of Clinical Isolates of *Salmonella enterica*. Applied and Environmental Microbiology **69**:4352–4358.

62. **Davidson CJ**, **White AP**, **Surette MG**. 2008. Evolutionary loss of the rdar morphotype in *Salmonella* as a result of high mutation rates during laboratory passage. ISME J **2**:293–307.

63. **Gerstel U**, **Park C**, **Romling U**. 2003. Complex regulation of *csgD* promoter activity by global regulatory proteins. Molecular Microbiology **49**:639–654.

64. **Gerstel U**, **Kolb A**, **Romling U**. 2006. Regulatory components at the *csgD* promoter - additional roles for OmpR and integration host factor and role of the 5' untranslated region. FEMS Microbiology Letters **261**:109–117.

65. **Ryjenkov DA**, **Simm R**, **Römling U**, **Gomelsky M**. 2006. The PilZ Domain Is a Receptor for the Second Messenger c-di-GMP. J Biol Chem **281**:30310–30314.

66. **García B**, **Latasa C**, **Solano C**, **Portillo FG-D**, **Gamazo C**, **Lasa I**. 2004. Role of the GGDEF protein family in *Salmonella* cellulose biosynthesis and biofilm formation. Molecular Microbiology **54**:264–277.

67. **Simm R**, **Lusch A**, **Kader A**, **Andersson M**, **Römling U**. 2007. Role of EAL-containing proteins in multicellular behavior of *Salmonella enterica* serovar Typhimurium. Journal of Bacteriology **189**:3613–3623.

68. **Solano C**, **García B**, **Latasa C**, **Toledo-Arana A**, **Zorraquino V**, **Valle J**, **Casals J**, **Pedroso E**, **Lasa I**, **Greenberg EP**. 2009. Genetic Reductionist Approach for Dissecting Individual Roles of GGDEF Proteins within the c-di-GMP Signaling Network in *Salmonella*. Proceedings of the National Academy of Sciences of the United States of America **106**:7997–8002.

69. **Mills E**, **Petersen E**, **Kulasekara BR**, **Miller SI**. 2015. A direct screen for c-di-GMP modulators reveals a *Salmonella* Typhimurium periplasmic L-arginine-sensing pathway. Science Signaling **8**:ra57–ra57.

70. **Simm R**, **Morr M**, **Kader A**, **Nimtz M**, **Römling U**. 2004. GGDEF and EAL domains inversely regulate cyclic di-GMP levels and transition from sessility to motility. Molecular Microbiology **53**:1123–1134.

71. **Yennamalli RM**, **Rader AJ**, **Kenny AJ**, **Wolt JD**, **Sen TZ**. 2013. Endoglucanases: insights into thermostability for biofuel applications. Biotechnol Biofuels **6**:136.

72. **Römling U**, **Galperin MY**. 2015. Bacterial cellulose biosynthesis: diversity of operons, subunits, products, and functions. Trends Microbiol **23**:545–557.

73. **Romling U**, **Galperin MY**, **Gomelsky M**. 2013. Cyclic di-GMP: the First 25 Years of a Universal Bacterial Second Messenger. Microbiology and Molecular Biology Reviews **77**:1–52.

74. **Simm R**, **Remminghorst U**, **Ahmad I**, **Zakikhany K**, **Romling U**. 2009. A role for the EAL-like protein STM1344 in regulation of CsgD expression and motility in *Salmonella enterica* serovar Typhimurium. Journal of Bacteriology **191**:3928.

75. **Kader A**, **Simm R**, **Gerstel U**, **Morr M**, **Römling U**. 2006. Hierarchical involvement of various GGDEF domain proteins in rdar morphotype development of *Salmonella enterica* serovar Typhimurium. Molecular Microbiology **60**:602–616.

76. **Ahmad I**, **Wigren E**, **Le Guyon S**, **Vekkeli S**, **Blanka A**, **Mouali el Y**, **Anwar N**, **Chuah ML**, **Lünsdorf H**, **Frank R**, **Rhen M**, **Liang Z-X**, **Lindqvist Y**, **Römling U**. 2013. The EAL-like protein STM1697 regulates virulence phenotypes, motility and biofilm formation in *Salmonella typhimurium*. Molecular Microbiology **90**:1216–1232.

77. **Da Re S**, **Ghigo J-M**. 2006. A CsgD-independent pathway for cellulose production and biofilm formation in *Escherichia coli*. Journal of Bacteriology **188**:3073–3087.

78. **Pontes MH**, **Lee E-J**, **Choi J**, **Groisman EA**. 2015. *Salmonella* promotes virulence by repressing cellulose production. Proceedings of the National Academy of Sciences **112**:5183–5188.

79. **Solomon EB**, **Niemira BA**, **Sapers GM**, **Annous BA**. 2005. Biofilm formation, cellulose production, and curli biosynthesis by *Salmonella* originating from produce, animal, and clinical sources. J Food Prot **68**:906–912.

80. **Holt KE**, **Thomson NR**, **Wain J**, **Langridge GC**, **Hasan R**, **Bhutta ZA**, **Quail MA**, **Norbertczak H**, **Walker D**, **Simmonds M**, **White B**, **Bason N**, **Mungall K**, **Dougan G**, **Parkhill J**. 2009. Pseudogene accumulation in the evolutionary histories of *Salmonella enterica* serovars Paratyphi A and Typhi. BMC Genomics **10**:36–12.

81. **Klemm EJ**, **Gkrania-Klotsas E**, **Hadfield J**, **Forbester JL**, **Harris SR**, **Hale C**, **Heath JN**, **Wileman T**, **Clare S**, **Kane L**, **Goulding D**, **Otto TD**, **Kay S**, **Doffinger R**, **Cooke FJ**, **Carmichael A**, **Lever AML**, **Parkhill J**, **MacLennan CA**, **Kumararatne D**, **Dougan G**, **Kingsley RA**. 2016. Emergence of host-adapted *Salmonella* Enteritidis through rapid evolution in an immunocompromised host. Nature Microbiology 1–6.

82. **Anriany YA**, **Weiner RM**, **Johnson JA**, **De Rezende CE**, **Joseph SW**. 2001. *Salmonella enterica* Serovar Typhimurium DT104 Displays a Rugose Phenotype. Applied and environmental microbiology **67**:4048–4056.

83. **Apel D**, **White A**, **Grassl G**, **Finlay B**, **Surette M**. 2009. Long Term Survival of *Salmonella enterica* Serovar Typhimurium reveals an infectious state that is underrepresented on laboratory media containing bile salts. Applied and Environmental Microbiology.

84. **Gonzalez-Escobedo G**, **Gunn JS**. 2013. Identification of *Salmonella enterica* Serovar Typhimurium Genes Regulated during Biofilm Formation on Cholesterol Gallstone Surfaces. Infection and Immunity **81**:3770–3780.

85. **Marshall JM**, **Flechtner AD**, **La Perle KM**, **Gunn JS**. 2014. Visualization of Extracellular Matrix Components within Sectioned *Salmonella* Biofilms on the Surface of Human Gallstones. PloS one **9**:e89243–9.

86. **Ogasawara H**, **Yamada K**, **Kori A**, **Yamamoto K**, **Ishihama A**. 2010. Regulation of the *Escherichia coli csgD* promoter: interplay between five transcription factors. Microbiology **156**:2470–2483.

87. **Steenackers H**, **Hermans K**, **Vanderleyden J**, **De Keersmaecker SCJ**. 2011. *Salmonella* biofilms: An overview on occurrence, structure, regulation and eradication. FRIN 1–30.

88. **Poppe, C.** 2000. *Salmonella* infections in the domestic fowl, p. 107–132. In C. Wray and A. Wray (ed.), *Salmonella* in domestic animals. Oxford University

Press, New York, NY.

6.0 GENERAL CONCLUSIONS, DISCUSSION, AND FUTURE DIRECTIONS

6.1 General Conclusions

- We have modified an existing Tn7 transposition system and demonstrated its improved efficiency for integrating genetic elements into the chromosome of *Salmonella* and other *Enterobacteriaceae*
- Plasmid-based complementation of a *csgD* mutant resulted in overexpression of both *csgD* and the biofilm phenotype
- Chromosomal-based complementation of a *csgD* mutant strain closely matched wildtype *Salmonella* in terms of timing of biofilm gene expression and quality of the biofilm phenotype
- Transcriptome analysis via RNA-seq identified 1856 genes that were differentially expressed between multicellular aggregates and planktonic cells exposed to the same environmental conditions
- The genetic profile of multicellular aggregates is consistent with previous characterizations of gene expression associated with the *Salmonella* rdar morphotype
- Planktonic cells display a genetic profile consistent with increased expression of virulence factors, including the SPI-1 and SPI-2 type three secretion systems, cell motility, and chemotaxis systems
- Planktonic cells synthesize a secretion-competent SPI-1 type three secretion system, which correlated with enhanced invasion of a polarized Caco-2 human intestinal cell line and a fitness advantage in the systemic model of murine infection compared to multicellular aggregates
- Multicellular aggregates demonstrated enhanced cell survival compared to planktonic cells following desiccation
- The fitness advantage of planktonic cells in the systemic model of murine infection was abolished when both cell subpopulations were pre-exposed to conditions of desiccation

- CsgD bistability may represent a conserved evolutionary strategy for generating phenotypic heterogeneity within the *Salmonella* population
- Planktonic cells and multicellular aggregates are specialized cell types that may enhance the transmission of *Salmonella*
- The invasive nontyphoidal *Salmonella* strains *S.* Typhimurium D23580 and *S.* Enteritidis D7795 are impaired for the biofilm phenotype
- Impairment of the biofilm phenotype of *S.* Typhimurium D23580 and *S.* Enteritidis D7795 are due to strain-specific genetic polymorphisms that affect the expression and/or function of genes associated with biofilm formation
- A *csgD* promoter mutation in *S.* Enteritidis D7795 correlated with loss of CsgD synthesis, absence of curli fimbriae and cellulose biosynthesis, and a biofilm-negative phenotype
- CsgD is synthesized by *S.* Typhimurium D23580, which correlated with detection of curli fimbriae and cellulose biosynthesis and an intermediate biofilm phenotype

6.2 General Discussion and Future Directions

The starting goal of my Ph.D. research was to characterize the transcriptome of *Salmonella* cells during the process of biofilm formation. Previously, *Salmonella* biofilm biology had been studied as a population-level phenomenon termed the rdar morphotype, where wildtype *Salmonella* colonies formed a highly organized extracellular matrix structure (1). CsgD, the master transcriptional regulator of the rdar morphotype, was known to activate genes that were responsible for the extracellular matrix structure itself (2). However, there was evidence that many other genetic changes occurred during biofilm formation as part of a global shift in the gene expression profile of *Salmonella* cells (3, 4). To determine the role of *csgD* in influencing the transcriptome during biofilm development, our approach was to compare biofilm-positive colonies of our wildtype *Salmonella* strain to biofilm-negative colonies of a *csgD* mutant strain. As a control for this and future experiments, we developed a *csgD* revertant strain using Tn7 transposition-based biotechnology and chromosomal complementation (Objectives 1 and 2 for this thesis, presented in Chapter 3).

Chromosomal complementation and Tn7-based transposition offered several attractive benefits for gene expression analysis. With plasmid-based complementation there was a concern regarding gene dosage effects; the biological mechanism of Tn7 transposition would instead allow for single-copy insertion of the gene of interest into an innocuous site in the chromosome. We chose the dual-plasmid Tn7 transposition system created by Choi *et al.* as it was heavily cited within the literature, and its suicide-vector approach allowed for generation of a chromosomal integrant in a single step (5). We found that a major limitation of this approach was the transient expression of the transposition machinery. While this issue was not implicitly stated in the literature, the authors had published sequential papers where new helper plasmid derivatives (pTNS2 and pTNS3) had been generated to improve expression of the transposition machinery via stronger promoter constructs (5, 6). Nancy Craig and colleagues offered a different solution by placing the transposition machinery and the mini-Tn7 transposon element onto a single plasmid with temperature-sensitive replication (7). Unfortunately, in our hands, cloning into the large plasmid was difficult, and when our construct was eventually cloned successfully, the procedure did not work (i.e., 20 independent experiments performed by four difference people). Our solution was to combine features of both systems by transferring the *tnsABCD* genes from pTNS2 onto a temperature-sensitive plasmid, while maintaining the original carrier vector for easier cloning (8). Our approach resulted in improved transposition efficiencies compared to the original dual-plasmid Tn7 transposition system when tested with pathogenic strains of *E. coli* and *Salmonella*. Further, we provided strong evidence that transposition success was correlated to increased expression of the *tnsABCD* genes from our new helper plasmid. In the 6 months following publication of our method, our new helper plasmid system has been requested by 15 laboratories around the world and has been successfully used by two of our collaborators for their chromosomal integration projects.

Our Tn7 transposition system provided a robust pipeline for chromosomal complementation. We evaluated plasmid- and chromosomal-complemented *csgD* revertant strains for their ability to restore the genetic and phenotypic biofilm profiles of the wildtype *S.* Typhimurium strain (8). Plasmid-based complementation resulted in overexpression of the *csgD* gene; this was demonstrated by the strain's overproduction of

curli fimbriae and cellulose and elevated *csgB* and *adrA* promoter expression compared to the wildtype strain. In contrast, the chromosomal-complemented strain was comparable to the wildtype strain both in the timing of gene expression and in the qualitative features of the biofilm phenotype. As it turned out, we did not use the *csgD* chromosomal revertant strain in the final transcriptomic study due to changes in the experimental design following the discovery of CsgD bistability. However, our research did make use of a second modification to the Tn7 system – the incorporation of the pCS26 reporter vector into the miniTn7 carrier plasmid. This allowed us to tag our wildtype *S.* Typhimurium strain with two different antibiotic markers for use in subsequent experiments, such as comparing the fitness of multicellular aggregates and planktonic cells in *in vitro* survival assays and in competitive infections of mice (Objective 8 of this thesis, presented in Chapter 4) (9). Therefore, our modified Tn7 transposition system allowed us to carry out important functional studies that validated and strengthened the results of the transcriptomic analysis.

To perform a successful RNA-seq experiment with *Salmonella* biofilms (Chapter 4), we were forced to develop several specialized methodologies for the genetic and physiological analysis of *Salmonella* biofilm formation. The reality is that most microbiological procedures are designed for the study of homogeneous populations of planktonic bacterial cells. We were faced with two main challenges: first, to modify existing tools to allow for reproducible extraction of biofilm cells embedded within the recalcitrant extracellular matrix, and second, to devise a way to compare the disparate cell physiologies of multicellular aggregates and planktonic cells. Our use of low speed centrifugation to isolate aggregated cells from planktonic cells was central to this study and provided an efficient method for separating the cell types with minimum impact to their transcriptome or physiology (Objective 4, presented in Chapter 4). For transcriptome analysis, we chose a column-based RNA extraction kit that allowed us to maximize our yield of high-quality RNA samples while facilitating the recovery of total RNA, including the small RNAs involved in *csgD* regulation. We successfully adapted a commercially available kit for the isolation of RNA from *Salmonella* biofilm samples (Objective 5, presented in Chapter 4). Our adjustments to minimize the effect of the extracellular matrix for cell lysis and on-column RNA purification resulted in a procedure

that elevated the power of our transcriptomic analysis due to its technical reproducibility. To compare the cell subpopulations by *in vitro* and *in vivo* functional assays, we generated additional procedures to accurately predict *Salmonella* cell numbers within multicellular aggregates. Our exhaustive development of a ratio of cell number to biofilm mass (Objective 6, presented in Chapter 4; Figure 4.3) allowed for us to accurately deliver equal quantities of cells from both subpopulations and enabled our fitness studies that compared these cell types in conditions of survival and infection. One issue we were faced with was to try to maintain the physical nature of the multicellular aggregates for these functional experiments, since previous studies had revealed that the extracellular matrix could impact gene expression (3). Initially, the adhesive nature of the extracellular matrix resulted in significant loss of cell aggregates that stuck to the surface of common laboratory tools such as pipette tips and microcentrifuge tubes. We observed that multicellular aggregates remain as individual, suspended "units" during growth in the culture media. Therefore, we used conditioned media as a suspension reagent in our experiments. This simple strategy abrogated the adhesive nature of the biofilm matrix and facilitated our studies that demanded the accurate delivery of a desired amount of biofilm cells. Altogether, our efforts resulted in the development of procedures and tools that made it possible to study the *Salmonella* cell subpopulations formed due to bistable CsgD expression and can be used by other researchers to study bacterial biofilms

In our main RNA-seq experiment, we characterized the gene expression profile of *Salmonella* multicellular aggregates and planktonic cells. Under the same environmental conditions, these clonal cell types had differential expression of over 1856 genes. The largest difference in gene expression between these subpopulations coincided with peak expression of *csgD* and the formation of multicellular aggregates. Previous work had found that several genes corresponding to carbon central metabolism and the cell stress response were expressed during biofilm formation (3). Our transcriptome analysis with RNA-seq expanded this finding by identifying the increased expression of genes important for the metabolism of amino acids, lipids, and nucleotides. This result supports our previous hypothesis that a metabolic shift accompanies the process of biofilm formation (3). The planktonic cell subpopulation and the CsgD-OFF state had never previously been characterized in the literature. We found that planktonic cells had a

comparable number of genes with increased expression as the multicellular aggregates. However, the transcriptomic profile of planktonic cells was vastly different from multicellular aggregates, with significant expression of multiple virulence traits, including the SPI-1 type three secretion system. In the literature, the dogma was that expression of the SPI-1 type three secretion system was exclusively induced by the host environment, which is an important first step in *Salmonella* pathogenesis. As such, SPI-1 expression in our experimental conditions was highly unexpected and required rigorous functional validation (Objectives 7 and 8, presented in Chapter 4). We confirmed that the proteins for the secretion apparatus and its effectors are synthesized under these atypical environmental conditions. Further, we showed that their synthesis provided a virulence advantage for planktonic cells compared to multicellular aggregates both in *in vitro* Caco-2 invasion assays and during competitive infections in susceptible mice. An important question remains of how expression of the SPI-1 type three secretion system is induced in the planktonic cell subpopulation. We found that several of the transcriptional activators associated with SPI-1 were expressed in planktonic cells, indicating that this inducing signal likely occurs early in the regulatory pathway. The relative absence of SPI-1 expression in multicellular aggregates may also provide an important clue for the molecular link between the persistence and virulence phenotypes. While it is tempting to speculate that there is a direct link between CsgD and SPI-1 expression, such a relationship has never before been described. It is plausible that the presence of the extracellular matrix may provide an important feedback signal that ultimately inhibits the expression of virulence factors such as the SPI-1 type three secretion system. Regulation between SPI-1 and biofilm formation may also be indirect in nature. Desai and colleagues recently demonstrated that SsrB, a transcriptional regulator encoded within SPI-2, can switch between promoting expression of the SPI-2 type three secretion system within the acidic macrophage vacuole and relieving H-NS silencing of *csgD* expression (10). Establishing the link between persistence and virulence is an important direction for the future of *Salmonella* biofilm research.

Our understanding of the *Salmonella* biofilm phenotype has historically been shaped by its characterization in the laboratory. The *in vitro* conditions that induced biofilm formation, together with the characteristics of the biofilm phenotype, suggested

that this physiology was important for the survival of *Salmonella* in non-host environments. An environmental setting likely presents many challenges to bacterial growth and replication, including nutrient limitation, extreme temperatures, pH fluctuations, UV and oxidative stress, and predation (11, 12). Therefore, the prevailing hypothesis was that *Salmonella* biofilm formation was an important mechanism for enhancing the persistence of cells in the non-host environment and improving the odds of their transmission to the next host (13). However, bistable CsgD expression and analysis of the cell subpopulations has shifted our perception of *Salmonella* biofilm formation and the process of pathogen transmission. What once was considered a population-level phenotype is now understood as a regulatory phenomenon mediated at the single cell level (14). Our research has uncovered an important, intimate link between the contrasting phenotypes of persistence and virulence. What is the purpose of this phenotypic heterogeneity in the life cycle of *Salmonella*? What is the advantage of expressing virulence factors in a non-host setting? Compared to the host niche, where the conditions of host-pathogen interactions are relatively defined, life in non-host environments and the process of transmission are unpredictable. Based on our characterization of the planktonic cell subpopulation, we have put forth the hypothesis that bistable CsgD expression is central to a bet-hedging strategy that improves the overall chances for *Salmonella* transmission. In a scenario where *Salmonella* immediately encounters the next host, planktonic cells would be readily able to instigate a new infection. In contrast, *Salmonella* biofilm cells would be prepared for extended exposure to the non-host environment prior to interactions with their next host. This unpredictable step in the *Salmonella* life cycle places equal selection pressure on virulence and persistence phenotypes. Through bistable CsgD expression, *Salmonella* cells can maintain the same genetic repertoire but simultaneously produce two disparate phenotypes.

Passage into non-host environments likely requires anticipatory genetic regulation, where *Salmonella* cells pre-emptively express the phenotype that is necessary for the next step in their life cycle. Thus, conditions within the gastrointestinal (GI) tract may also provide important cues for bistable CsgD expression. In the GI niche, *Salmonella* cells are exposed to host temperatures, low pH, high osmolarity, bile acids,

antimicrobial peptides, iron limitation, and in some cases, nutrient limitation (12, 15). Some of these conditions (i.e. temperature and osmolarity) may favour the expression of virulence factors such as the SPI-1 type three secretion system (15), while repressing *Salmonella* biofilm formation (1, 16). Conversely, exposure to stresses such as bile, antimicrobial peptides, iron limitation, or poor nutrient availability would promote *csgD* expression and the biofilm phenotype. Regulation of *csgD* expression and synthesis is incredibly complex, and few studies have attempted to understand the hierarchy of this regulation (1, 16-18). For example, iron limitation has been shown to over-ride the normal temperature shut-off of biofilm formation at 37˚C (1). It is possible that microenvironments within the intestinal niche provide strong activating signals required to generate the CsgD-ON state. In the lumen of the small intestine, high bile concentrations have been shown to increase the intracellular concentration of cyclic-di-GMP in the enteropathogen *Vibrio cholerae* (reviewed in (12)). For *Salmonella*, high concentrations of c-di-GMP activate *csgD* expression and promote the biofilm phenotype (19). Similarly, Diard *et al.* recently demonstrated that expression of the SPI-1 type three secretion system is location-dependent within the intestine (20). *Salmonella* cells positioned at the surface of the intestinal epithelial layer were 100% positive for SPI-1 type three secretion expression, while the majority of cells in the lumen were negative for this virulence trait. We predict that in anticipation of exit from the host, biofilm expression may also be induced by signals produced towards the end of host-pathogen interactions, including intestinal inflammation and the associated rapid replication of *Salmonella* within the lumen (21-24). The reactive oxygen species produced by neutrophils undoubtedly provides an important cue for activation of the bacterial cell stress response. Metabolic cues, such as the pathways involved with anaerobic respiration during this stage (i.e., ethanolamine and 1,2-propanediol) or the nutrient limitation caused by rapid *Salmonella* replication following inflammation, may also add a temporal element to the anticipatory regulation of *Salmonella* biofilm formation. Consistent with this, our transcriptomic analysis revealed that genes for the ethanolamine and 1,2-propanediol metabolic pathways have increased expression within multicellular aggregates (9).

Our ability to elucidate the *in situ* regulation of *Salmonella* biofilm formation is limited by the reality that *Salmonella* biofilms have yet to be observed in nature. However, the enteric bacteria *Vibrio cholerae* has been shown to exist both as planktonic cells and in multicellular aggregates in human stool samples (25, 26). Previously, White *et al.* investigated *csgD* expression and the biofilm phenotype during murine infection by using *Salmonella* with luciferase reporters fused to the biofilm-related gene promoters (13). They observed that *csgD* expression was activated within the mouse intestine during the course of infection, but that expression of curli biosynthesis genes (*csgBAC*) was only observed within fecal pellets that had passed out of the infected mice (13). Based on this result, they concluded that synthesis of the extracellular matrix occurs after passage from the host. However, work by Dylan Shivak as part of his M.Sc. thesis demonstrated that there is a limitation in the ability to detect of bacterial luciferase during the course of infection (8, 27). In his work, Shivak found that heavy bacterial loads were required for luciferase detection, despite expression from a strong constitutive promoter (8). If activation of extracellular matrix production were limited to only a subpopulation of *Salmonella* cells, signaling from the *csgBAC* promoter would not be detectable using bioluminescence as a marker. Work in our laboratory is currently underway to determine if curli fimbriae and multicellular aggregates can be observed during murine infections by using single cells detection techniques (i.e., tissue sectioning and confocal microscopy).

Our work with invasive nontyphoidal *Salmonella* strains from sub-Saharan Africa follows from the results of our RNA-seq experiment. As explained in the Literature Review chapter, invasive nontyphoidal *Salmonella* are a phylogenetically distinct group of nontyphoidal *Salmonella* that are responsible for invasive bloodstream infections (28, 29). Like typhoidal serovars, strains of invasive nontyphoidal *Salmonella* induce a fever-like illness and persist within the systemic compartment of infected individuals (30). At the current time, humans are the only identified reservoir for invasive strains of nontyphoidal *Salmonella* (31, 32). There is an established correlation that biofilm formation is highly conserved in nontyphoidal *Salmonella* serovars that can infect a broad host range and cause gastroenteritis in humans (33-36). In contrast, typhoidal *Salmonella* serovars associated with systemic infections and a restricted host range have lost the capacity for biofilm formation (33). At the time of our study, the biofilm phenotype of

invasive nontyphoidal *Salmonella* strains had never been determined. However, the significant mortality rate associated with invasive nontyphoidal *Salmonella* infections (an estimated 681 316 deaths annually) emphasized the need to better understand the transmission of invasive nontyphoidal *Salmonella* transmission. We hypothesized that what we learned about *Salmonella* biofilm formation could help us to understand more about the lifecycle of invasive nontyphoidal *Salmonella*. Based on their intimate association with the human host, we predicted that invasive nontyphoidal *Salmonella* strains would be negative for the biofilm phenotype.

We focused our evaluation of invasive nontyphoidal *Salmonella* on *S.* Typhimurium D23580 and *S.* Enteritidis D7795, two strains with recently completed genome sequences. These strains were compared to a panel of biofilm-positive, nontyphoidal strains and biofilm-negative, typhoidal *Salmonella* strains. Genome sequences from all panel strains were obtained from databases or were completed by us; we wondered if we would be able to demonstrate a consistent pattern of gene degradation in *Salmonella* biofilm formation that is shared between invasive nontyphoidal and typhoidal *Salmonella* strains. We know that host-adapted and host-restricted *Salmonella* strains, including Typhi and invasive nontyphoidal *Salmonella*, demonstrate a consistent pattern of genomic degradation (29, 37), where the genes associated with intestinal replication following inflammation are functionally inactive (24). These genes are functional in nontyphoidal *Salmonella* serovars associated with gastroenteritis and are thought to play an essential role in their transmission, allowing for an intestinal bloom and subsequent shedding of high bacterial loads from the host (38). Phenotypically, both invasive nontyphoidal strains were impaired in their biofilm-forming capacity; *S.* Enteritidis D7795 was negative in every assay, whereas *S.* Typhimurium D23580 had an intermediate phenotype (i.e., partial biofilm pattern on its colony surface, the slime-like texture of its multicellular aggregates, and the detection of trace amounts of curli fimbriae and cellulose polymers). Our observations with *S.* Typhimurium D23580 did not match the first published report stating that this strain was biofilm-negative (37). We assume that the differences in our results could reflect disparities in the strain histories and perhaps whether they had been passaged in vitro (i.e., our version was not). *S.* Enteritidis D7795 and all typhoidal *Salmonella* strains included in our strain panel were negative for

curli and cellulose biosynthesis, and the CsgD protein could not be detected. We identified an inactivating *csgD* promoter mutation in *S.* Enteritidis D7795, while strains of *S.* Typhi were mutated within the coding sequence for *csgD*. *S.* Typhimurium D23580 also had two *cis*-acting mutations upstream of the *csgD* promoter, which correlated with a significant reduction in *csgD* transcriptional expression. However, CsgD protein levels could be detected in lysates obtained from this strain. We assume that the biofilm-negative phenotype of *S.* Enteritidis D7795 can be fully explained by the csgD promoter mutation, whereas for *S.* Typhimurium D23580 the answer is more complex. Our preliminary genomic comparisons (which are incomplete and thus not reported in my thesis) identified a single nucleotide polymorphism in STM1987 (STM14_2408), which encodes an important cyclic-di-GMP producing enzyme that positively regulates biofilm formation. This polymorphism exists within the sensory domain of the STM1987 protein that is associated with activating the enzyme in response to sensing an unknown signal from the external environment. Genome engineering experiments are currently underway in the White laboratory to determine if these identified mutations in *S.* Enteritidis D7795 and *S.* Typhimurium D23580 can be linked to the impairment in biofilm formation. Our study represents a thorough analysis of biofilm formation in two invasive nontyphoidal *Salmonella* strains and should form the benchmark for the future analysis of any additional invasive nontyphoidal strains.

My body of work has identified a new aspect of *Salmonella* biology that has important implications for understanding the transmission of this pathogen. Nontyphoidal *Salmonella* strains that briefly colonize the host and cause gastroenteritis must contend with survival in both the host and non-host environment. Bistable expression of *csgD*, biofilm formation, and specialized subpopulations of multicellular aggregates and planktonic cells represent an evolutionary adaptation for mitigating the unpredictable nature of transmission via the fecal-oral route. In contrast, invasive strains of nontyphoidal and typhoidal *Salmonella* have evolved to persist chronically within the systemic niche of the host. This feature of host adaptation relieves the selection pressure placed on *Salmonella* to persist within an environmental reservoir and correlates with loss of the biofilm phenotype. Identifying *Salmonella* reservoirs is crucial for developing effective public health efforts to reduce disease burden. Recent advancements in DNA

sequencing technology have vastly expanded the availability of pathogen genomic sequences, and have allowed unprecedented characterization of both classical and emerging *Salmonella* strains. Invasive nontyphoidal *Salmonella* strains are an example of how genome sequencing can be complemented by our understanding of *Salmonella* pathogenesis to develop a genetic signature for host adaptation and predict the nature of host-pathogen interactions. Similarly, we predict that our gene expression profile of *Salmonella* biofilm formation could be used to predict the transmission properties of *Salmonella* strains. Future research efforts should use our transcriptomic analysis as a guide to isolate the core genetic processes that govern biofilm formation. This core biofilm genetic signature would be an invaluable aid for predicting the nature of transmission of *Salmonella* strains associated with future outbreaks.

6.3 References

1. **Romling U**, **Sierralta WD**, **Eriksson K**, **Normark S**. 1998. Multicellular and aggregative behaviour of *Salmonella typhimurium* strains is controlled by mutations in the agfD promoter. Molecular Microbiology **28**:249–264.

2. **Brombacher E**, **Baratto A**, **Dorel C**, **Landini P**. 2006. Gene expression regulation by the Curli activator CsgD protein: modulation of cellulose biosynthesis and control of negative determinants for microbial adhesion. Journal of Bacteriology **188**:2027–2037.

3. **White A**, **Weljie A**, **Apel D**, **Zhang P**. 2010. A Global Metabolic Shift Is Linked to *Salmonella* Multicellular Development. PloS one. **5**:e11814.

4. **Hamilton S**, **Bongaerts RJ**, **Mulholland F**, **Cochrane B**, **Porter J**, **Lucchini S**, **Lappin-Scott HM**, **Hinton JC**. 2009. The transcriptional programme of *Salmonella enterica* serovar Typhimurium reveals a key role for tryptophan metabolism in biofilms. BMC Genomics **10**:599–21.

5. **Choi K-H**, **Gaynor JB**, **White KG**, **Lopez C**, **Bosio CM**, **Karkhoff-Schweizer RR**, **Schweizer HP**. 2005. A Tn7-based broad-range bacterial cloning and expression system. Nat Meth **2**:443–448.

6. **Choi KH**, **Mima T**, **Casart Y**, **Rholl D**, **Kumar A**, **Beacham IR**, **Schweizer HP**. 2008. Genetic Tools for Select-Agent-Compliant Manipulation of *Burkholderia pseudomallei*. Applied and Environmental Microbiology **74**:1064–1075.

7. **McKenzie GJ**, **Craig NL**. 2006. Fast, easy and efficient: site-specific insertion of transgenes into Enterobacterial chromosomes using Tn*7* without need for selection

of the insertion event. BMC Microbiology **6**:39.

8. **Shivak DJ**, **MacKenzie KD**, **Watson N**, **Pasternak A**, **Jones BD**, **Wang Y**, **DeVinney R**, **Wilson H**, **Surette MG**, **White AP**. 2016. A modular, Tn7-based system for making bioluminescent or fluorescent *Salmonella* and *E. coli* strains. Applied and Environmental Microbiology AEM.01346–16–15.

9. **MacKenzie KD**, **Wang Y**, **Shivak DJ**, **Wong CS**, **Hoffman LJL**, **Lam S**, **Kröger C**, **Cameron ADS**, **Townsend HGG**, **Köster W**, **White AP**. 2015. Bistable Expression of CsgD in *Salmonella enterica* Serovar Typhimurium Connects Virulence to Persistence. Infection and Immunity **83**:2312–2326.

10. **Desai SK**, **Winardhi RS**, **Periasamy S**, **Dykas MM**, **Jie Y**, **Kenney LJ**. 2016. The horizontally-acquired response regulator SsrB drives a *Salmonella* lifestyle switch by relieving biofilm silencing. eLife **5**:e10747.

11. **Winfield MD**, **Groisman EA**. 2003. Role of Nonhost Environments in the Lifestyles of *Salmonella* and *Escherichia coli*. Applied and Environmental Microbiology **69**:3687–3694.

12. **Silva AJ**, **Benitez JA**. 2016. Vibrio cholerae Biofilms and Cholera Pathogenesis. PLoS Negl Trop Dis **10**:e0004330–25.

13. **White A**, **Gibson D**, **Grassl G**, **Kay W**, **Finlay B**, **Vallance B**, **Surette M**. 2008. Aggregation via the red, dry, and rough morphotype is not a virulence adaptation in *Salmonella* enterica serovar Typhimurium. Infection and Immunity **76**:1048.

14. **Grantcharova N**, **Peters V**, **Monteiro C**, **Zakikhany K**, **Romling U**. 2010. Bistable expression of CsgD in biofilm development of *Salmonella enterica* serovar Typhimurium. Journal of Bacteriology **192**:456.

15. **Altier C**. 2005. Genetic and environmental control of *Salmonella* invasion. J Microbiol **43 Spec No**:85–92.

16. **Gerstel U**, **Park C**, **Romling U**. 2003. Complex regulation of *csgD* promoter activity by global regulatory proteins. Molecular Microbiology **49**:639–654.

17. **Gerstel U**, **Romling U**. 2001. Oxygen tension and nutrient starvation are major signals that regulate agfD promoter activity and expression of the multicellular morphotype in *Salmonella typhimurium*. Environmental Microbiology **3**:638–648.

18. **Kader A**, **Simm R**, **Gerstel U**, **Morr M**, **Römling U**. 2006. Hierarchical involvement of various GGDEF domain proteins in rdar morphotype development of *Salmonella enterica* serovar Typhimurium. Molecular Microbiology **60**:602–616.

19. **Romling U**, **Galperin MY**, **Gomelsky M**. 2013. Cyclic di-GMP: the First 25 Years of a Universal Bacterial Second Messenger. Microbiology and Molecular

Biology Reviews **77**:1–52.

20. **Diard M**, **Garcia V**, **Maier L**, **Remus-Emsermann MNP**, **Regoes RR**, **Ackermann M**, **Hardt W-D**. 2014. Stabilization of cooperative virulence by the expression of an avirulent phenotype. Nature **494**:353–356.

21. **Thiennimitr P**, **Winter SE**, **Winter MG**, **Xavier MN**, **Tolstikov V**, **Huseby DL**, **Sterzenbach T**, **Tsolis RM**, **Roth JR**, **Bäumler AJ**. 2011. Intestinal inflammation allows *Salmonella* to use ethanolamine to compete with the microbiota. Proceedings of the National Academy of Sciences of the United States of America **108**:17480–17485.

22. **Thiennimitr P**, **Winter SE**, **Bäumler AJ**. 2012. *Salmonella*, the host and its microbiota. Current Opinion in Microbiology **15**:108–114.

23. **Winter SE**, **Thiennimitr P**, **Winter MG**, **Butler BP**, **Huseby DL**, **Crawford RW**, **Russell JM**, **Bevins CL**, **Adams LG**, **Tsolis RM**, **Roth JR**, **Bäumler AJ**. 2010. Gut inflammation provides a respiratory electron acceptor for *Salmonella*. Nature **467**:426–429.

24. **Nuccio SP**, **Baumler AJ**. 2014. Comparative Analysis of *Salmonella* Genomes Identifies a Metabolic Network for Escalating Growth in the Inflamed Gut. mBio **5**:e00929–14–e00929–14.

25. **Faruque SM**, **Biswas K**, **Udden SMN**, **Ahmad QS**, **Sack DA**, **Nair GB**, **Mekalanos JJ**. 2006. Transmissibility of cholera: in vivo-formed biofilms and their relationship to infectivity and persistence in the environment. Proceedings of the National Academy of Sciences **103**:6350–6355.

26. **Nelson EJ**, **Chowdhury A**, **Harris JB**, **Begum YA**, **Chowdhury F**, **Khan AI**, **LaRocque RC**, **Bishop AL**, **Ryan ET**, **Camilli A**, **Qadri F**, **Calderwood SB**. 2007. Complexity of rice-water stool from patients with *Vibrio cholerae* plays a role in the transmission of infectious diarrhea. PNAS **104**:19091–19096.

27. **Shivak DJ**. 2015. Exploring the Dynamics of *Salmonella* Transmission in a Murine Model of Infection. University of Saskatchewan.

28. **Kingsley RA**, **Msefula CL**, **Thomson NR**, **Kariuki S**, **Holt KE**, **Gordon MA**, **Harris D**, **Clarke L**, **Whitehead S**, **Sangal V**, **Marsh K**, **Achtman M**, **Molyneux ME**, **Cormican M**, **Parkhill J**, **MacLennan CA**, **Heyderman RS**, **Dougan G**. 2009. Epidemic multiple drug resistant *Salmonella* Typhimurium causing invasive disease in sub-Saharan Africa have a distinct genotype. Genome Research **19**:2279–2287.

29. **Hadfield J**, **Keddy KH**, **Dallman TJ**, **Jacobs J**, **Deng X**, **Wigley P**, **Barquist LB**, **Langridge GC**, **Feltwell T**, **Harris SR**, **Mather AE**, **Fookes M**, **Aslett M**, **Msefula C**, **Kariuki S**, **MacLennan CA**, **Onsare RS**, **Weill FCO-X**, **Le Hello S**, **Smith AM**, **McClelland M**, **Desai P**, **Parry CM**, **Cheesbrough J**, **French N**,

Campos J, **Chabalgoity JA**, **Betancor L**, **Hopkins KL**, **Nair S**, **Humphrey TJ**, **Lunguya O**, **Cogan TA**, **Tapia MD**, **Sow SO**, **Tennant SM**, **Bornstein K**, **Levine MM**, **Lacharme-Lora L**, **Everett DB**, **Kingsley RA**, **Parkhill J**, **Heyderman RS**, **Dougan G**, **Gordon MA**, **Thomson NR**, **Feasey NA**. 2016. Distinct *Salmonella* Enteritidis lineages associated with enterocolitis in high-income settings and invasive disease in low-income settings. Nature Genetics 1–9.

30. **Feasey NA**, **Dougan G**, **Kingsley RA**, **Heyderman RS**, **Gordon MA MD**. 2012. Invasive non-typhoidal *Salmonella* disease: an emerging and neglected tropical disease in Africa. The Lancet **379**:2489–2499.

31. **Kariuki S**, **Revathi G**, **Kariuki N**, **Kiiru J**, **Mwituria J**, **Muyodi J**, **Githinji JW**, **Kagendo D**, **Munyalo A**, **Hart CA**. 2006. Invasive multidrug-resistant non-typhoidal *Salmonella* infections in Africa: zoonotic or anthroponotic transmission? Journal of medical microbiology **55**:585–591.

32. **Kariuki S**, **Revathi G**, **Gakuya F**, **Yamo V**, **Muyodi J**, **Hart CA**. 2002. Lack of clonal relationship between non-typhi *Salmonella* strain types from humans and those isolated from animals living in close contact. FEMS Immunol Med Microbiol **33**:165–171.

33. **Römling U**, **Bokranz W**, **Rabsch W**, **Zogaj X**, **Nimtz M**, **Tschäpe H**. 2003. Occurrence and regulation of the multicellular morphotype in *Salmonella* serovars important in human disease. International Journal of Medical Microbiology **293**:273–285.

34. **Solomon EB**, **Niemira BA**, **Sapers GM**, **Annous BA**. 2005. Biofilm formation, cellulose production, and curli biosynthesis by *Salmonella* originating from produce, animal, and clinical sources. J Food Prot **68**:906–912.

35. **Solano C**, **García B**, **Valle J**, **Berasain C**, **Ghigo J-M**, **Gamazo C**, **Lasa I**. 2002. Genetic analysis of *Salmonella enteritidis* biofilm formation: critical role of cellulose. Molecular Microbiology **43**:793–808.

36. **White A**, **Surette M**. 2006. Comparative genetics of the rdar morphotype in *Salmonella*. Journal of Bacteriology **188**:8395.

37. **Singletary LA**, **Karlinsey JE**, **Libby SJ**, **Mooney JP**, **Lokken KL**, **Tsolis RM**, **Byndloss MX**, **Hirao LA**, **Gaulke CA**, **Crawford RW**, **Dandekar S**, **Kingsley RA**, **Msefula CL**, **Heyderman RS**, **Fang FC**. 2016. Loss of Multicellular Behavior in Epidemic African Nontyphoidal *Salmonella* enterica Serovar Typhimurium ST313 Strain D23580. mBio **7**:e02265–15–11.

38. **Lawley TD**, **Bouley DM**, **Hoy YE**, **Gerke C**, **Relman DA**, **Monack DM**. 2008. Host transmission of *Salmonella enterica* serovar Typhimurium is controlled by virulence factors and indigenous intestinal microbiota. Infection and Immunity **76**:403–416.